13273-8

NOVEMBER 1989

ACADEMIC MEDICINE

[illegible]

Legal Context for Dismissals

Teaching and Learning in the Ambulatory Setting

Appleton-Lange Medical Student Essay Contest: Ethical Challenges to Physicians in the 1990s

Commentaries: Rx: Writing, The Physician-Patient Dyad

Independent Study in Medical Science

American-Soviet Medical Exchange Program

Assessments of Patients' Health in Rural Areas

A Health Promotion and Disease Prevention Curriculum

CME Programs and Physicians' Transfusion Behavior

Standardizing Dean's Letters

AAMC Data Report: Teaching Hospitals

VOLUME 64 • NUMBER ELEVEN

ASSISTANT OR ASSOCIATE PROFESSOR
(EDUCATIONAL SPECIALIST)

Requirements: Ph.D. in Education, 3-4 years' Postdoctoral experience, demonstrated record of publications.

Responsibilities: 1) Plan, implement, and evaluate research in education of Medical Students on the Surgical service as well as of Residents. 2) Serve as consultant for educational matters to the Chairman of the departmental curriculum committee as well as the departmental Chairman. 3) Plan, implement, and evaluate ways to improve teaching effectiveness of the surgical faculty. 4) Provide consultation to the third year curriculum Director for improvement of the third year curriculum for Medical Students.

Please send curriculum vitae to: **James W. Mackenzie, Professor and Chairman, Department of Surgery, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, One Robert Wood Johnson Place, CN-19, (JME), New Brunswick, NJ 08903-0019.**

The UMDNJ is an Affirmative Action/Equal Employment Opportunity Employer M/F/H/V.

UMD
NEW JERSEY

ADVERTISING IN

ACADEMIC MEDICINE

The AAMC encourages display advertising for all academic and administrative posts in the journal, which publishes at the beginning of each month. Deadlines are two months before publication, but every effort will be made to accommodate ads received as late as one month prior to publication. "Camera-ready" copy is preferable, but the publisher will typeset and lay out copy at an additional charge.

For advertising rates and other information, contact:

Diane Sherel, (215) 546-0313

Pass or Fail?

A National Conference ...

Examining Medical Education and Current Reforms

January 26, 27 & 28, 1990

Stanford University School of Medicine

Sponsored by the Stanford Health Policy Forum

...Featuring...

Dr. Gordon Moore: Founder of Harvard's New Pathways Program; Director of Teaching Programs

Dr. Howard Barrows: Dean of Educational Affairs; Southern Illinois School of Medicine

Dr. Arthur Kaufman: Program Director for Family Medicine; UNM Medical School

Dr. Kelley Skeff: Housestaff Program Director, Faculty Development; Stanford Medical School

...Addressing...

Socialization of Physicians

Minority Issues in Medical Education

Student and Resident Hours

Curriculum Reform

Student Evaluation

Clinical Teaching

Problem-Based Learning

CME Credits Available

Registration: SHPF, Health Research and Policy Bldg., Rm. 7, Stanford University, Stanford, CA 94305 (415) 725-2243

Springer publishing co

Clinical Problem-Based Learning
Workbook for Integrated Basic and Clinical Science
R.E. Waterman et al., Editors
This book provides students with techniques for developing clinical reasoning skills and integrating basic science information with everyday patient problems. Illustrated 1988/ 322 pp $19.95

Innovation in Medical Education
An Evaluation of its Present Status
Zohair M. Nooman, Henk G. Schmidt, and Esmat S. Ezzat, Editors
Bringing together the experiences of 30 medical schools, this book provides data comparing the effects of innovative versus conventional medical curricula. Also discussed are the possible effects of changes recommended by the GPEP report. Dec. 1989/ 448 pp (est)

Successful Faculty in Academic Medicine
Essential Skills and How to Acquire Them
C.J. Bland et al., Eds. Identifies and describes skills physicians need to succeed as faculty and proposes detailed plans for teaching these skills through a model faculty development curriculum. 1989/ 336 pp/ $38.95.

Springer Publishing Company
536 Broadway • New York • NY • 10012 H16

Letters to the Editor 667

Research Reports

ARTICLES

Patients' and Physicians' Assessments of Patients' Health in Rural Mississippi *by Paul C. Lampley, Ph.D., and Bruce Randolph,* M.D., M.P.H. 669

A First-Year Problem-based Curriculum in Health Promotion and Disease Prevention *by William C. Taylor,* M.D., *Richard J. Pels,* M.D., *and Robert S. Lawrence,* M.D. 673

Assessing University Nephrology Training as Preparation for Community Consultative Practice *by Richard S. Muther,* M.D., *Marilyn Murray,* R.N., *and William M. Bennett,* M.D. 677

Effects of a CME Program on Physicians' Transfusion Practices *by Alan L. Hull, Ph.D., Jay Wasman,* M.D., *and Lawrence T. Goodnough,* M.D. 681

Evaluating Preclinical Medical Students by Using Computer-based Problem-Solving Examiantions by Ronald H. Stevens, Ph.D., Anthony R. Kwak, Ed.D., and J. Michael McCoy, M.D. 685

SHORT PAPER

A New Approach to Standardizing the Dean's Letter *by Norma E. Wagoner, Ph.D., and J. Robert Suriano, Ph.D.* 688

BRIEF COMMUNICATIONS

A Comparison of Students in Six-year and Traditional Eight-Year Medical School Programs on Measures of Personality and Stress Early in Medical School *by Ronald G. Nathan, Ph.D., Frances E. Nixon,* M.D., *Lee Bairnsfather, Ph.D., John H. Allen, Ph.D., and Mary Hack,* J.D. 690

Cross-Cultural Validation of the Doctor-Patient Scale *by Ronald J. Markert, Ph.D.* 690

Use of Class Ranking in Deans' Letters *by Clinton H. Toewe II,* M.D., *and David R. Golay,* M.U.R.P. 690

Analysis of Student Time Allocation on a Medicine Clerkship *by Ruth-Marie E. Fincher,* M.D., *Lloyd A. Lewis, Ph.D., and Leonard D. Nance,* III, M.S. 691

AAMC Data Report

Characteristics of COTH Member, Other Teaching, and Non-teaching Hospitals, 1987 *by Joyce V. Kelly, Ph.D.* 692

Bibliography 694

PUBLISHING SERVICES: **Hanley & Belfus, Inc.** 10 South 13th Street, Philadelphia, PA 19107

NOVEMBER 1989
VOLUME 64
NUMBER ELEVEN

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Innovations in Medical Education
Wayne K. Davis, Ph.D.
University of Michigan Medical School

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, M.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman M. Mellinkoff, M.D.
University of California, Los Angeles
School of Medicine

CONSULTING EDITORS
L. Thompson Bowles, M.D., Ph.D.
George Washington University
School of Medicine and Health Sciences
Milton Corn, M.D.
Georgetown University School of Medicine
Robert M. Heyssel, M.D.
The Johns Hopkins Health System
Donald G. Kassebaum, M.D.
Association of American Medical Colleges
Russell Miller, M.D.
Howard University
Jay P. Sanford, M.D.
Uniformed Services University of the Health Sciences
F. Edward Hébert School of Medicine
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center
School of Medicine

Chairman-Elect
David H. Cohen, Ph.D.
Northwestern University

Immediate Past Chairman
John W. Colloton, M.A.
University of Iowa Hospitals and Clinics

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University
School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Ernst R. Jaffé, M.D.
Albert Einstein College of Medicine

Chairman, Council of Teaching Hospitals
Gary Gambuti
St. Luke's-Roosevelt Hospital Center

EDITORIAL BOARD

L. Thompson Bowles, M.D., Ph.D. (Chairman)
George Washington University
School of Medicine and Health Sciences
Philip C. Anderson, M.D.
University of Missouri, Columbia
School of Medicine
G. William Bates, M.D.
Medical University of South Carolina
College of Medicine
Nancy E. Gary, M.D.
Albany Medical College
David S. Greer, M.D.
Brown University Program in Medicine
Paul F. Griner, M.D.
Strong Memorial Hospital
Rochester, New York
Kaaren I. Hoffman, Ph.D.
University of Southern California
School of Medicine
John E. Ives
St. Luke's Episcopal Hospital
Houston, Texas
Fernando S. Mendoza, M.D.
Stanford University School of Medicine
Gordon S. Page, Ed.D
University of British Columbia
Faculty of Medicine
Hugh M. Scott, M.D.
Bishop's University
Lennoxville, Quebec
Charles E. Spooner, Jr., Ph.D.
University of California, San Diego
School of Medicine
Manuel Tzagournis, M.D.
Ohio State University College of Medicine

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, N.W., Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1989 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.

Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

DATA REPORT

Characteristics of COTH Member, Other Teaching, and Non-teaching Hospitals, 1987

Teaching hospitals,* which provide settings for the clinical education of physicians, nurses, and most other health care workers, comprise 22% of all short-term general non-federal hospitals (hereafter referred to as hospitals) in the country. Those major teaching hospitals that are members of the Council of Teaching Hospitals (COTH)† represent 6% of all hospitals and 27% of all teaching hospitals.

Urban or rural location and number of beds are the predominant structural factors distinguishing teaching and non-teaching hospitals. Virtually all COTH members (99%) and nearly all other teaching hospitals‡ (91%) are located in urban areas, while only 42% of non-teaching hospitals are urban. The number of hospital beds also is closely associated with teaching status. In terms of average bed size, COTH members have 548 beds, other teaching hospitals have 288 beds, and non-teaching facilities have 114 beds. Thus, the clinical education of future practitioners occurs almost exclusively in large urban hospitals.

Although they represent only 6% of all hospitals, COTH members account for 21% of all hospital beds, 23% of hospital admissions, and 28% of hospital outpatient visits. All teaching hospitals together account for 22% of hospitals and approximately half of hospital beds, admissions, and outpatient visits. Hospital occupancy rates also are associated with teaching status: on average, 77% for COTH members, 65% for other teaching hospitals, and 51% for non-teaching hospitals. However, there is wide variation among non-teaching hospitals, with small rural hospitals having the lowest rates.

COTH hospitals are the locus of clinical education for 79% of housestaff (interns and residents), while 19% train in other teaching facilities. Eighty-five percent of other health care workers receive their training in teaching hospitals (data not shown). The large number of residents (on average, 136 full-time-equivalent positions for COTH members) contribute a substantial quantity of medical services to patients hospitalized in major teaching hospitals.

COTH member hospitals provide a disproportionate share of care to Medicaid and charity patients. While accounting for 23% of total hospital admissions, COTH members admit 34% of all Medicaid patients and provide 59% of all charity care, measured by the dollar value of uncollectible charges. COTH hospitals provide a greater share of unreimbursed care for several reasons: their urban locations, the presence of large outpatient departments, the inclusion of many large urban municipal hospitals among COTH members, and hospital policies of accepting low-income and uninsured patients.

The average cost per adjusted admission (the number of

Table 1

Selected Operating Characteristics of Teaching and Non-teaching Hospitals, 1987*

Characteristic	Teaching Hospitals: COTH Members	Teaching Hospitals: Others	Non-teaching Hospitals	All
Number	358	927	4,337	5,622
Percent of all hospitals	6	16	77	100
Percent urban	99	91	42	53
Average number of beds	548	288	114	170
Percent of all beds	21	28	52	100
Percent of all admissions	23	29	48	100
Percent of all hospital outpatient visits	28	27	45	100
Percent occupancy	77	65	51	55
Percent of all housestaff	79	19	1	100
Average number of housestaff (FTE)	136	13	0	11
Percent of all Medicaid admissions	34	27	39	100
Percent of all charity care	59	25	16	100
Average cost per adjusted admission ($)	5,511	4,195	3,075	3,415

*The study population was all short-term general non-federal hospitals. For definition of other terms, see the text and text footnotes. Source for data: American Hospital Association Annual Survey Data, 1987.

---

*Teaching hospitals are defined as institutions participating in one or more accredited graduate medical education (GME or medical residency) programs that provide supervised clinical training for recent medical school graduates. Accreditation is provided by the Accreditation Council for Graduate Medical Education.

†To qualify as teaching hospital members of COTH, hospitals must sponsor or significantly participate in at least four active accredited residency programs. At least two of these programs must be in medicine, surgery, obstetrics-gynecology, pediatrics, family practice, or psychiatry. Exceptions to the requirement for four residency programs may be made for specialty institutions.

‡Other teaching hospitals are those non-COTH institutions participating in one or more accredited GME programs.

admissions adjusted to include volume of outpatient activity) also is related to teaching status. Other teaching hospitals have an average cost per adjusted admission that is 36% higher than costs in non-teaching hospitals ($4,195 compared with $3,075). For COTH members, costs are 31% higher than those in other teaching hospitals ($5,511 compared with $4,195). Investigators have not determined precisely why teaching hospital costs are higher and posit a number of determining factors: treating patients who are more severely ill or have more complicated illnesses; higher wage rates and other expenses associated with an urban location; providing more intensive treatment, including more tests, ancillary services, and nursing care; acquiring and operating advanced technology; maintaining stand-by capacity associated with a referral center; and inefficiencies associated with educating residents.

In addition to their patient care and educational missions, teaching hospitals are the primary sites of clinical research, which accounts for additional operating costs that are incompletely funded. The 123 academic medical centers§ represented in COTH play a major role in the development and dissemination of biomedical and health services research. These research efforts require the combination of expert staff and sophisticated technology that characterizes major medical centers.

Joyce V. Kelly, Ph.D.

Dr. Kelly is associate vice president, Division of Clinical Services, Association of American Medical Colleges, Washington, D.C.

These analyses were conducted with support provided by The Commonwealth Fund and data provided by the Hospital Data Center of the American Hospital Association.

§Academic medical center hospitals are defined as those COTH members under common ownership with a college of medicine, or having the majority of medical school department chairmen as the hospital chiefs of service, or having the chairmen responsible for appointing the hospital chiefs of service.

A3273-8

DECEMBER 1989

ACADEMIC MEDICINE

Teaching Medical Ethics

Literature Review

Medical Education: Coming of Age

Program Descriptions

- Baylor
- Chicago
- Loyola
- Brown-RI
- Northwestern
- East Carolina
- Penn State
- Pittsburgh
- UCSF and Washington

Research

Book Reviews: Ethics Topics

VOLUME 64 • NUMBER TWELVE

DECEMBER 1989
VOLUME 64
NUMBER TWELVE

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Foreword

Special Issue: Teaching Medical Ethics *by Addeane S. Caelleigh* 699

Preface *by Laura Weiss Lane*, M.D., *and Steven H. Miles*, M.D. 700

Overviews

Teaching Medical Ethics: Some Persistent Questions and Some Responses *by Edmund D. Pellegrino*, M.D. 701

Teaching Medical Ethics: A Hard-won Beachhead *by Edward J. Stemmler*, M.D. 704

Literature Review

Medical Ethics Education: Coming of Age *by Steven H. Miles*, M.D., *Laura Weiss Lane*, M.D., *Janet Bickel*, M.A., *Robert M. Walker*, M.D., *and Christine K. Cassel*, M.D. 705

Program Descriptions

The Baylor Experience in Teaching Medical Ethics *by Baruch A. Brody*, Ph.D. 715

Medical Ethics Teaching Programs at the University of California, San Francisco, and the University of Washington *by Albert R. Jonsen*, Ph.D. 718

Development of a Teaching Program in Clinical Medical Ethics at the University of Chicago *by Robert M. Walker*, M.D., *Laura Weiss Lane*, M.D., *and Mark Siegler*, M.D. 723

Development of the Medical Humanities Program at East Carolina University *by Loretta M. Kopelman*, Ph.D. 730

The Clinical Medical Humanities Program at Loyola University of Chicago *by David C. Thomasma*, Ph.D., *and Patricia Marshall*, Ph.D. 735

Ethics Education at Northwestern University Medical School *by James F. Bresnahan*, S.J., J.D., LL.M., Ph.D., *and Kathryn Montgomery Hunter*, Ph.D. 740

Teaching Medical Ethics in Its Contexts: Penn State College of Medicine *by David Barnard*, Ph.D., *and K. Danner Clouser*, Ph.D. 744

Evolution of Clinical Ethics Teaching at the University of Pittsburgh *by Joel Frader,* M.D., *Robert Arnold,* M.D., *John Coulehan,* M.D., *Rosa Lynn Pinkus,* Ph.D., *Alan Meisel,* J.D., *and Kenneth Schaffner,* Ph.D., M.D. 747

The Development of Medical Ethics Curriculum in a General Internal Medicine Residency Program *by Steven A. Wartman, M.D., Ph.D., and Dan W. Brock, Ph.D.* 751

Research Reports
ARTICLES

The Effect of Teaching Medical Ethics on Medical Students' Moral Reasoning *by Donnie J. Self, Ph.D., Fredric D. Wolinsky, Ph.D., and DeWitt C. Baldwin, Jr., M.D.* 755

Internal Medicine Residents' Preferences Regarding Medical Ethics Education *by Jay A. Jacobson,* M.D., *Susan W. Tolle,* M.D., *Carol Stocking,* Ph.D., *and Mark Siegler,* M.D. 760

Book Reviews

Essays on the Rise and Decline of Bedside Medicine (Altschule) *Reviewed by* C. Rollins Hanlon, M.D. 765
Racial Hygiene (Proctor) *Reviewed by* Fritz Redlich, M.D. 766
The AIDS Patient: An Action Agenda (Rogers and Ginzberg) *Reviewed by* Philip R. Lee, M.D. 767
House Officer: Becoming a Medical Specialist (Cohen) *Reviewed by* Louis J. West, M.D. 768
Cases in Bioethics: Selections from the Hastings Center Report (Levine) *Reviewed by* Leslie Steven Rothenberg, J.D. 769
Technology, Bureaucracy and Healing in America: A Post-Modern Paradigm (Bulger) *Reviewed by* David M. Rogers, M.D. 770
Disease in the Popular American Press: The Case of Diphtheria, Typhoid Fever, and Syphilis, 1870–1920 (Ziporyn) *Reviewed by* Phil Manning, M.D. 772

Innovations in Medical Education

Evaluating Clinical Teachers for Promotion *by Arthur I. Rothman, Ed.D., Peeter Poldre, M.D., and Robert Cohen, Ph.D.* 774

From the President 776

AAMC FOCUS 777

National Policy Perspectives 778

Letters to the Editor 780

AAMC Data Report 782

Bibliography 783

The Index to Volume 64 will appear in the January 1990 issue

DECEMBER 1989
VOLUME 64
NUMBER TWELVE

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Innovations in Medical Education
Wayne K. Davis, Ph.D.
University of Michigan Medical School

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, M.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman M. Mellinkoff, M.D.
University of California, Los Angeles
School of Medicine

CONSULTING EDITORS
L. Thompson Bowles, M.D., Ph.D.
George Washington University
School of Medicine and Health Sciences
Milton Corn, M.D.
Georgetown University School of Medicine
Robert M. Heyssel, M.D.
The Johns Hopkins Health System
Donald G. Kassebaum, M.D.
Association of American Medical Colleges
Russell Miller, M.D.
Howard University
Jay P. Sanford, M.D.
Uniformed Services University of the Health Sciences
F. Edward Hébert School of Medicine
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center
School of Medicine

Chairman-Elect
David H. Cohen, Ph.D.
Northwestern University

Immediate Past Chairman
John W. Colloton, M.A.
University of Iowa Hospitals and Clinics

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University
School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Ernst R. Jaffé, M.D.
Albert Einstein College of Medicine

Chairman, Council of Teaching Hospitals
Gary Gambuti
St. Luke's-Roosevelt Hospital Center

EDITORIAL BOARD

L. Thompson Bowles, M.D., Ph.D. (Chairman)
George Washington University
School of Medicine and Health Sciences
Philip C. Anderson, M.D.
University of Missouri, Columbia
School of Medicine
G. William Bates, M.D.
Medical University of South Carolina
College of Medicine
Nancy E. Gary, M.D.
Albany Medical College
David S. Greer, M.D.
Brown University Program in Medicine
Paul F. Griner, M.D.
Strong Memorial Hospital
Rochester, New York
Kaaren I. Hoffman, Ph.D.
University of Southern California
School of Medicine
John E. Ives
St. Luke's Episcopal Hospital
Houston, Texas
Fernando S. Mendoza, M.D.
Stanford University School of Medicine
Gordon S. Page, Ed.D
University of British Columbia
Faculty of Medicine
Hugh M. Scott, M.D.
Bishop's University
Lennoxville, Quebec
Charles E. Spooner, Jr., Ph.D.
University of California, San Diego
School of Medicine
Manuel Tzagournis, M.D.
Ohio State University College of Medicine

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, N.W., Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1989 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.

Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

IV

Internal Medicine Residents' Preferences Regarding Medical Ethics Education

JAY A. JACOBSON, M.D., SUSAN W. TOLLE, M.D., CAROL STOCKING, Ph.D., and MARK SIEGLER, M.D.

Abstract—Three hundred and twenty-three residents in six internal medicine programs in three states were surveyed concerning what they wanted to learn about medical ethics and how they would prefer to learn it; they were also asked to indicate what medical ethics education they had already received. Specifically, the residents were given a list of 35 medical ethics topics and asked to indicate whether they would like more attention to a topic, or whether it had received enough or too much attention. (They could also indicate that they thought the topic was inappropriate for attention during residency.) The residents were also given a list of 17 teaching methods for medical ethics and asked to indicate whether each method was very useful, somewhat useful, or not useful. Sixty-one percent of the residents responded. Approximately three-fourths of these had had some formal teaching about medical ethics in both medical school and residency, and nearly all indicated that they wanted more ethics training on specific topics. Certain topics were chosen as particularly suitable or unsuitable for such training; many of those judged as deserving more attention concerned legal issues and end-of-life issues. Most of the responding residents regarded standard clinical teaching formats as very useful for learning about medical ethics but regarded several other methods as not useful. The authors discuss the implications of these and related findings. *Acad. Med.* 64(1989):760–764.

Internal medicine residency is an opportune and appropriate time to teach medical ethics,[1–6] but just what to teach to residents and how to teach it remain perplexing questions. Striking variations in the residents' prior experience with ethics,[7,8] new issues arising from rapid changes in medicine, and compelling clinical demands on their time make it essential to choose topics that residents regard as important and methods that are efficient and effective.

To discover what residents want to learn about medical ethics and how they would prefer to learn it, we conducted a questionnaire survey of residents in six training programs in three states. Our principal objective was to provide program directors and ethics educators with a list of subject areas that most residents wanted to learn more about and methods of teaching that residents indicated would be most effective.

Method

We surveyed residents from six internal medicine programs: The University of Chicago and the Evanston Hospital in Illinois, the University of Utah and the LDS Hospital in Salt Lake City, Utah, and the Oregon Health Sciences University and the Good Samaritan Hospital in Portland, Oregon. We selected three university programs and three private affiliated programs. The program directors or the chief residents notified the house staff about our survey and encouraged them to participate. In February 1989, we distributed a six-page questionnaire about medical ethics education to the 332 residents in these programs. A second mailing to nonrespondents was sent in April 1989.

The 35 medical ethics topics we asked the residents about were similar to those used in a previous survey of physicians and those mentioned in the DeCamp report.[3,5] The residents we surveyed could indicate whether they would like more attention to a topic, or whether it had received enough or too much attention. They could also indicate that they thought the topic was inappropriate for attention during residency. The 17 teaching methods we offered for consideration were those discussed in the DeCamp Report or described by other educators.[9–13] The residents we surveyed were asked to indicate whether each method was very useful, somewhat useful, or not useful. Residents were also asked to indicate what medical ethics education they had received in medical school and during their residency periods.

We assured residents that their responses would be confidential and that only aggregated data would be reported. The design of the study and the questionnaire were approved by the Institutional Review Board at the University of Chicago Pritzker School of Medicine.

Results

Of the 332 residents, 202 (61%) returned their questionnaires by May 12, 1989. The sizes of their programs ranged from 33 to 77 residents, and

---

Dr. Jacobson and Dr. Tolle were visiting scholars, Center for Clinical Medical Ethics, Pritzker School of Medicine, University of Chicago, Chicago, Illinois, at the time of the study. Dr. Jacobson is chief, Division of Medical Ethics, LDS Hospital and University of Utah Department of Medicine in Salt Lake City, Utah; Dr. Tolle is director, Center for Ethics in Health Care, Oregon Health Sciences University, Portland, Oregon. Dr. Stocking is director of research and Dr. Siegler is director, Center for Clinical Medical Ethics, Pritzker School of Medicine.

Correspondence and requests for reprints should be addressed to Dr. Jacobson, Division of Medical Ethics, LDS Hospital, 8th Avenue and C St., Salt Lake City, UT 84143.

the response rates varied by program from 52 to 75%. Overall, 65% of the sample were men and 62% of the respondents were men. The average age of the residents was 30 years. They were approximately equally divided in terms of years of training: 37% were in their first year, 30% in their second year, and 30% in their third year. In two programs, approximately 60% of the responses came from first-year trainees. The 162 residents who identified their medical schools had graduated from 79 U.S. medical schools.

Seventy-two percent of the residents had had at least one formal course about medical ethics in medical school, usually during the preclinical years; 97% recalled some informal education about ethics as well. A majority in each program and 74% overall said they had received medical ethics teaching during their residencies (range, 57–90%). There was no significant difference in the answers to this question by the residents' years of training.

The residents' preferences about ethics topics and how to teach them are shown in Tables 1 and 2. Surprisingly, the responses regarding topics did not differ significantly between those who had had no formal ethics course in medical school and those who had had one or more. With two exceptions, the distribution of responses did not vary significantly by the residents' years of training. The first-year residents were more likely than their more senior colleagues to want more teaching about withholding life support (80% versus 64%, chi-square = 6.077, df = 2, $p = .05$), and the more senior residents more often wanted additional information about laws related to physician-patient relations (84% versus 71%, chi-square = 5.81, df = 2, $p = .055$).

Nineteen topics merit more attention, according to a majority or plurality of residents in the survey. No topic was judged to have received too much attention by more than 5% of respondents, and no topic was felt to be inappropriate by more than 13%.

Table 1 shows the 11 topics identified by 50% or more of the residents in each program as needing more attention. The list suggests that residents are particularly interested in learning more about laws and decisions related to end-of-life issues and about problems of access to health care and allocation of resources.

Table 1 also contains 17 topics selected by a majority of residents in at least one program, but not by all, as needing more attention. Distributions of the responses by program differed significantly for six topics: role of an ethics consultant, functions of ethics committees, conflict resolution between resident and attending physician and conflict resolution between doctors and patients, and the family's and the patients roles in decision making (chi − square > 31.7, df = 15, $p < .01$).

Only seven topics were judged not to need more attention by a majority of residents in any program. Only four topics were considered inappropriate by at least 10% of all respondents: making admission or treatment decisions based on ability to pay, experimentation on human subjects, religious views on ethical questions, and euthanasia.

The residents' responses about the usefulnesses of various teaching methods are shown in Table 2. Although there was variation between programs, several methods were fairly strongly endorsed or rejected. Four methods were judged very useful by a majority of residents: several conferences on ethics, several lectures or grand rounds, incorporation of these issues in other presentations and discussions, and discussion with colleagues and a knowledgeable discussion leader. Only two methods were regarded as not useful by more than 50%: watching a self-instructive videotape about an ethical issue and a reading course based on novels.

Discussion

Our survey showed that most of the residents had had some medical ethics education in medical school,[7,8] that 74% also had had it during their residencies, and that nearly all wanted more ethics training in specific areas. Most regarded standard clinical teaching formats as very useful for learning about medical ethics, but they regarded several other methods as not useful.

A majority of respondents from each program we surveyed indicated that they had had some formal teaching about ethics during their residencies. This high prevalence of ethics education may not be representative of all U.S. programs. In a 1984 survey, 40% of internal medicine residencies provided no formal exposure to medical ethics.[14] Despite some ethics education during residency, most of our respondents indicated that they would like more. These results parallel those of several earlier surveys of residents and practicing physicians.[15,16] Three conclusions, then, can be drawn from our survey.

First, residents, including those with prior exposure to teaching about medical ethics, appear receptive to more medical ethics education. However, how much time should be spent on medical ethics was not addressed in our survey.

Second, despite variations in their prior experiences and differences between their residency programs, a majority of respondents from each program chose 11 topics that they would like to learn more about. These clustered in the areas of laws and practices related to limiting life-support systems, allocation of resources, and access to medical care. Our respondents' interest in these subjects was keener than that of physicians in an earlier survey.[5] This may be because, as residents, they confront end-of-life decisions more often than many practicing physicians do and because physicians in general are more concerned today about economic issues and legal risk than they were in 1982. However, even topics not cited by most residents as meriting more attention might be of considerable interest and importance to individual residents because of their particular circumstances.

Third, some topics may be particularly well or poorly suited for individual programs. Several of the differences we encountered among programs regarding specific topics are

Table 1

Percentages of 202 Residents in Six Internal Medicine Residency Programs* Who Reported They Wanted More Attention Given to Instruction on Each of 37 Topics in Medical Ethics, 1989

Topic	% of All Respondents	Range of % by Program
Selected by 50% or more of residents in all programs		
Laws related to withdrawal of life support	84	71–95
Laws related to physician-patient relations	80	71–90
Withdrawing life-supporting treatment	75	58–88
Living wills and advance directives	73	61–95
Allocation of limited resources related to diagnosis or prognosis	73	50–86
Government policy regarding access to medical care	73	61–80
Withholding life-supporting treatment	70	52–80
Understanding rights of incompetent patients	64	53–74
Euthanasia	63	50–75
Incompetent or impaired colleagues	62	55–85
Priority for scarce organs	61	50–75
Selected by 50% or more of residents in one or more, but not all, programs		
Making admission or treatment decisions based on ability to pay	60	43–85
Selection of transplant recipients	56	34–72
Requesting organ donation	54	45–72
Do not rescitate orders	51	42–60
Withholding information from patients	49	34–63
Role of an ethics consultant†	49	16–85
Experimentation on human subjects	48	30–55
Religious views on ethical questions	45	26–80
How to care for and communicate with a dying patient	47	34–67
Functions of ethics committees†	46	26–80
How to cope with death of patient	44	36–57
Conflict resolution between resident and attending†	42	25–60
Conflict resolution between doctors and nurses	42	28–55
Conflict resolution between doctors and patients†	40	31–50
Nursing ethics	39	16–75
Confidentiality	38	22–54
Physician's conflict of interest: financial advantage versus patient welfare	34	11–50
Not selected by 50% or more of residents in any program		
Family's role in decision making†	32	10–46
Patient's role in decision making†	30	13–40
Informed consent	35	24–45
Physician's obligations to patients	34	22–45
Personal values versus patient's values	33	19–43
Guidance on expression of compassion	30	16–45
Physician's role in decision making	26	13–40

*The six programs (one in Chicago and one in Evanston, Illinois; two in Salt Lake City, Utah; and two in Portland, Oregon) had a total of 332 residents.

†Significant difference between programs, $p < .01$.

explainable and instructive. Residents at a program with a widely discussed policy regarding admission of indigent patients were less likely than others to want more attention paid to making admission or treatment decisions based on ability to pay.

Likewise, many house officers at a program with an active medical ethics consultation service desired less additional information about the role of an ethics consultant. Most residents at a program with a newly established ethics committee, which included house staff members, indicated that they wanted more education about the function of that committee. Thus, in a particular program, awareness of issues, of the resources available, and of the frequency with which particular problems arise should help predict which topics residents would like to learn more about. A survey similar to ours could be used by ~~an in~~dividual program to plan its curriculum.

While our survey records residents' preferences for learning about ethical problems, it does not describe the types of ethical problems physicians encounter or their frequencies. These are important considerations for curriculum planning, and other studies have examined them.[17–20] Physicians surveyed in 1982 reported that the four most common issues they confronted were compassion, truth-tell-

Table 2

Percentages of 202 Residents in Six Internal Medicine Residency Programs* Who Rated as "Very Useful" Each of 17 Methods for Teaching Medical Ethics, 1989

Teaching Method	% of All Respondents	Range of % by Program
Several conferences on ethics†	57	33–76
Incorporation of these issues in other presentations and discussions†	57	30–78
Several lectures or grand rounds on ethics†	55	24–83
Discuss these subjects with colleagues and knowledgeable discussion leader	53	36–59
Calling on ethics consultation†	45	6–63
Special lecture course on ethics†	42	12–67
Observing attending physicians handle ethical problems with patient	40	15–56
Bringing problems to the attention of an ethics committee†	37	6–60
Independent reading	34	27–47
Discuss ethical issues with colleagues	32	23–48
Discuss ethical issues with hospital attorney	30	20–35
Workshops at professional meetings	30	15–41
An ethics manual	29	9–45
Reading aided by annotated bibliography	29	24–36
Discussing a videotape about an ethical issue with knowledgeable discussion leader	19	9–25
A reading course based on novels	9	5–16
Watching self-instructive videotape	8	0–11

*The six programs (one in Chicago and one in Evanston, Illinois; two in Salt Lake City, Utah; and two in Portland, Oregon) had a total of 332 residents. They rated the teaching methods on a Likert-type scale.

†Significant differences between programs, $p < .01$.

ing, patient's role in decision making, and obtaining informed consent.[5] On a medical inpatient service, sensitized residents most often identified as problems withholding tests or treatment, informed consent, truth-telling, relationships among physicians, and limited resources.[1] All of these problems, except informed consent, were among those categorized most often in our survey as suitable for more attention.

Our survey provides some helpful information about methods for teaching ethics. The residents clearly rejected watching instructional videotapes by themselves or reading novels, but they were not asked specifically about other uses of videotapes or about reading shorter pieces of literature. The majority of residents favored the methods by which they learn about the rest of clinical medicine. This may be because they have little experience with some of the other methods included in our survey. We do not believe that the survey results argue convincingly against using these techniques, especially in view of the wide variation in responses by program. The results, however, suggest that conventional methods, available at nearly all programs, may be quite acceptable for medical ethics teaching. Other, innovative methods not listed in our survey may also be useful.[11]

Several internal medicine departments have already described the contents and methods they are using in their ethics programs[9-13] and others do so in this issue of *Academic Medicine*. Barnard has called attention to other, perhaps neglected dimensions of medical ethics.[21]

Although the response rate of residents to our survey was better than the rate achieved in similar surveys,[5] the number of nonrespondents, the lack of information about whether they differed significantly from respondents, and the non-random selection of programs preclude broad generalizations from our results. The age and gender distributions of our participating residents, for example, are similar to those of internal medicine residents in the United States, but our survey included no acknowledged foreign medical graduates. Nevertheless, we think the results provide some guidance, especially for programs similar to the ones we studied.

We hope our survey of residents' preferences combined with the observations, recommendations, and curricular innovations of others will help faculty provide the kind of medical ethics education that most residents seem eager to receive.

The authors thank the program directors, house staff coordinators, chief medical residents, and especially the residents themselves, for their cooperation and participation in this study; and Thelma Davis for helping in the conduct of the study. The work of the Center for Clinical Medical Ethics is supported by The Henry J. Kaiser Family Foundation, The Pew Charitable Trusts, and the Andrew W. Mellon Foundation.

References

1. Lo, B., Schroeder, M. D. Frequency of Ethical Dilemmas in a Medical Inpatient Service. *Arch. Intern. Med.* **141**(1981): 1062–1064.
2. Winkenwerder, W. Jr. Ethical Dilemmas for House Staff Physicians. *JAMA* **254**(1985):3454–3457.
3. Culver, C. M., et al. Basic Curricular Goals in Medical Ethics. *N. Engl. J. Med.* **312**(1985):253–256.
4. Subcommittee on the Evaluation of Humanistic Qualities in the Internist, American Board of Internal Medicine. Evaluation of Humanistic Qualities of the Internist. *Ann. Intern. Med.* **99**(1983): 720–724.
5. Pellegrino, E. D., Hart, R. J., Henderson, S. R., Loeb, S. E., and Edwards, G. Rele-

vance and Utility of Courses in Medical Ethics. *JAMA* **253**(1985):49–53.

6. Siegler, M. A Legacy of Osler: Teaching Clinical Ethics at the Bedside. *JAMA* **239**(1978):951–956.
7. Pellegrino, E. D., and McElhinney, T. K. Teaching Ethics, The Humanities and Human Values in Medical Schools: A Ten-Year Overview. Washington, D.C.: Institute on Human Values in Medicine and Society for Health and Human Values, 1982.
8. Bickel, J. Integrating Human Values Teaching Programs into Medical Students' Clinical Education. Washington, D.C.: Association of American Medical Colleges, 1986.
9. Radwany, S. M., and Adelson, B. H. The Use of Literary Classics in Teaching Medical Ethics to Physicians. *JAMA* **257**(1987):1629–1631.
10. Arnold, R. M., Forrow, L., Wartman, S. A., and Teno, J. Teaching Clinical Ethics: A Model Programme for Primary Care Residency. *J. Med. Ethics* **14**(1988):91–96.
11. Perkins, H. S. Teaching Medical Ethics in Internal Medicine Residency: A Rationale and a Proposal for Curricular Content and Format. *Arch. Intern. Med.* (in press).
12. Tolle, S. W., Cooney, T. G., and Hickam, D. H. A Program to Teach Residents Humanistic Skills for Notifying Survivors of a Patient's Death. *Acad. Med.* **64**(1989): 505–506.
13. Hong, H. P., Singer, P. A., Lynch, A., and Siegler, M. Teaching and Learning Clinical Medical Ethics during Residency Training. *Ann. R. Coll. Phys. Surg. Can.* **21**(1988): 2193–2198.
14. Povar, G. J., and Keith, K. J. The Teaching of Liberal Arts in Internal Medicine Residency Training. *J. Med. Ed.* **59**(1984): 714–721.
15. Kearn, D. C., Parrino, T. A., and Korst, D. R. The Lasting Value of Clinical Skills. *JAMA* **254**(1985):70–76.
16. Sheehan, T. J., et al. Teaching Humanistic Behavior. Presented at the annual meeting of the American Education Research Association, Washington, D.C., April 1987.
17. Connelly, J. D., and DalleMura, S. Ethical Problems in the Medical Office. *JAMA* **260**(1988):812–815.
18. La Puma, J. Stocking, C. B., Silverstein, M.D., DiMartini, A., and Siegler, M. An Ethics Consultation Service in a Teaching Hospital—Utilization and Evaluation. *JAMA* **260**(1988):808–811.
19. Perkins, H. S., and Saathoff, B. S. Impact of Medical Ethics Consultations on Physicians: An Exploratory Study. *Am. J. Med.* **85**(1988):761–765.
20. Brennan, T. Ethics Committees and Decisions to Limit Care, Experience at Massachusetts General Hospital. *JAMA* **260** (1988):803–807.
21. Barnard, D. Residency Ethics Teaching: A Critique of Current Trends. *Arch. Intern. Med.* **148**(1988):1836–1838.

AAMC DATA REPORT

Expectations of Medical School Expenses and Financing among 1988 MCAT Examinees

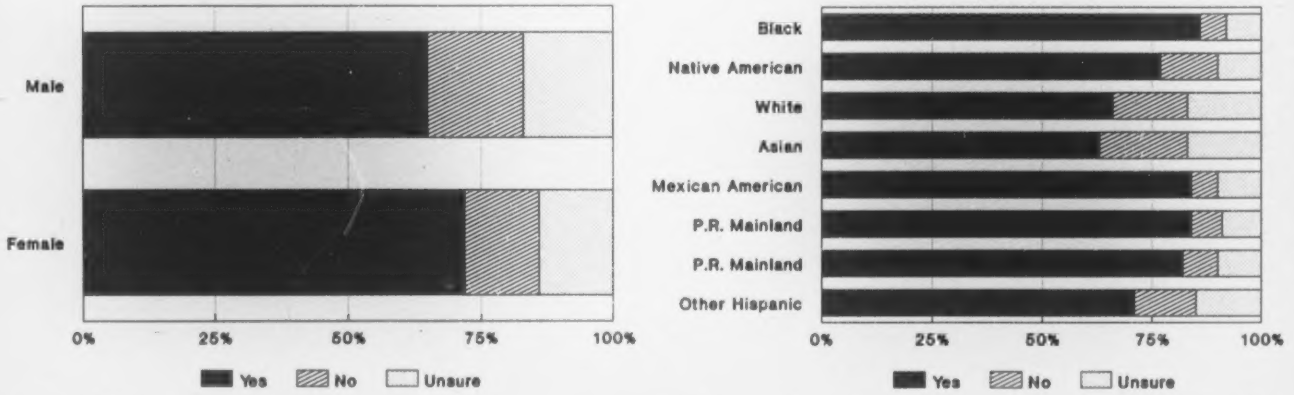


***Figure 1.* Percentages of 1988 MCAT Examinees who expected to need financial aid for their medical education, by gender (*left*) and by ethnicity (*right*). Sample was the 81% (26,944) who answered financial questions on the Premedical Student Questionnaire: by gender (men = 15,842, women = 11,102); by ethnicity (black = 2,139, Native American = 134, white, except Hispanic = 18,196, Asian or Pacific Islander = 3,662, Mexican American = 498, Puerto Rico Mainland = 214, Puerto Rico Commonwealth = 489, other Hispanic = 584, ethnic group not identified = 1,038.**

The figures and tables shown here present data from the Premedical Student Questionnaire (PMQ) completed by the 33,007 examinees who took the Medical College Admission Test (MCAT) in 1988. Approximately 81% of them (26,944) gave information on how they expected to pay for medical school.

When asked whether they anticipated needing financial assistance to complete medical school, 68% responded yes, 16% responded no, and 16% were unsure. Figure 1 shows that the percentages who anticipated needing aid varied substantially by gender and ethnicity. Overall, the amount they anticipated needing was $40,007. The average amount they expected to borrow varied substantially by ethnicity and region, less so by gender (Table 1).

The students' expectations are similar to the actual average indebtedness and distribution of debt among graduating medical students. As reported in the 1988 AAMC Medical Student Graduation Questionnaire, the average indebtedness of graduates with debt was $38,489, compared with the $40,007 anticipated by the 1988 PMQ respondents. The proportion of 1988 graduates without debt (17%) is close to the percentage of PMQ respondents indicating that they would not need financial aid (16%). Thus, premedical students seem to have fairly realistic expectations about the cost of medical education.

Wendy Luke Colquitt, M.A.

Ms. Colquitt is a research associate, Section for Operational Studies, Association of American Medical Colleges, Washington, D.C.

Table 1

Average Amounts of Loans that 1988 MCAT Examinees Anticipated They Would Need to Finance Their Medical Education, by Gender, Ethnicity, and Region*

Groups	Estimated Loan Needed	No. of Respondents
Gender		
Men	$39,219	9,212
Women	41,094	6,687
Ethnicity		
Black	41,599	1,604
Native American	44,446	91
White (except Hispanic)	39,562	10,395
Asian or Pacific Islander	40,209	2,024
Mexican-American	42,337	366
Puerto Rican (Mainland)	47,799	161
Puerto Rican (Commonwealth)	38,418	354
Other Hispanic	38,020	373
Region		
Northeast	47,390	3,781
South	35,907	4,550
Central	39,437	3,704
West	41,954	2,615
All respondents	$40,007	15,899

*Of the 33,007 who took the 1988 Medical College Admission Test, 81% (26,944) answered the financial questions on the Premedical Student Questionnaire, and 68% of those who answered expected to need financial aid; overall, 15,899 specified the amount.

B3273-8

JANUARY 1990

ACADEMIC MEDICINE

Job Turnover of Hospital CEOs

Evaluating Innovative Curricula

Computers in Academic Medicine

- Databases in the Basic Sciences
- A Computer Workstations Project
- Evaluating Instructional Software
- Hypermedia Technology

Teaching Agenda in Research Ethics

Evaluating Residents

Views of Educational Needs

Volunteer Community Preceptors

Practitioners' Career Changes

VOLUME 65 • NUMBER ONE

JANUARY 1990
VOLUME 65
NUMBER ONE

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Invited Article
Job Turnover of CEOs in Teaching and Nonteaching Hospitals, *by Peter A. Weil, Ph.D.* 1

Special Article
Charting the Winds of Change: Evaluating Innovative Medical Curricula, *by Charles P. Friedman, Ph.D., Ruth de Bliek, M.Ed., David S. Greer, M.D., Stewart P. Mennin, Ph.D., Geoffrey R. Norman, Ph.D., Cecil G. Sheps, M.D., David B. Swanson, Ph.D., and Christel A. Woodward, Ph.D.* 8

Ideas in Medical Education
Introduction, *by Janine C. Edwards, Ph.D* 15
Computer Databases as an Educational Tool in the Basic Sciences, *by Charles P. Friedman, Ph.D., Robert G. Twarog, Ph.D., Dean D. File, Patricia L. Youngblood, Ph.D., and Ruth de Bliek, M.Ed.* 15

Medical Informatics
The Case for Hypermedia, *by Mark E. Frisse, M.D.* 17

Academic Computing in Medicine
Introduction, *by Robert F. Rubeck, Ph.D.* 20
First Year's Experience of the MAClinical Computer Workstations Project, *by Thomas O. Stair, M.D., Milton Corn, M.D., and Naomi C. Broering, M.L.S., M.A.* 20
Evaluation by Second-Year Medical Students of Their Computer-Aided Instruction, *by George Xakellis, M.D., and Craig Gjerde, Ph.D.* 22

Foundations
The Henry J. Kaiser Family Foundation, *by Edward L. Schor, M.D.* 26

From the President 28

AAMC Focus 29

National Policy Perspectives 30

Commentary
Research Ethics: A Teaching Agenda for Academic Medicine, *by Paul J. Friedman*, M.D. 32

Editorial
Changes for 1990, *by Addeane S. Caelleigh* 34

Letters to the Editor
A National Physician Service Corps? *by Milford Fulop*, M.D. 34

Increasing Autopsies *by Rolla B. Hill*, M.D.; *Susan W. Tolle*, M.D., *Thomas G. Cooney*, M.D., *David H. Hickam*, M.D., M.P.H. 35

Research Reports
ARTICLES
A Methodological Framework for the Design of Research on the Evaluation of Residents, *by Donald A. Risucci*, Ph.D., *and Anthony J. Tortolani*, M.D. 36

Humanism in Medical Education: A Study of Educational Needs Perceived by Trainees of Three Canadian Schools, *by Brigitte Maheux*, M.D., Ph.D., *Pierre Delorme*, M.D., M.Sc., *François Béland*, Ph.D., *and Jean Beaudry*, M.A. 41

SHORT PAPERS
A Workshop Program to Train Volunteer Community Preceptors, *by Joseph M. Keenan*, M.D., *Harold C. Seim*, M.D., M.P.H., *Carole J. Bland*, Ph.D., *and Thomas M. Altemeier*, M.D. 46

Major Changes in Medical Careers Following Medical School Graduation: When, How Often, and Why, *by Penny A. Jennett*, Ph.D., *Michael Kishinevsky*, *Heather Bryant*, M.D., Ph.D., *and Kay L. Hunter*, M.Sc. 48

Bibliography 50

Index to Volume 64 55

Publishing Services: Hanley & Belfus, Inc.
210 South 13th Street, Philadelphia, PA 19107

JANUARY 1990
VOLUME 65
NUMBER ONE

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
  Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
  St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
  Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
  Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
  University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
  Arizona Health Sciences Center

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
  Georgetown University Hospital
Philip Anderson, M.D.
  University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
  Harvard Medical School
Eta Berner, Ed.D.
  University of Alabama School of Medicine
Ruth Bulger, Ph.D.
  National Academy of Sciences
Vincent A. Fulginiti, M.D.
  Tulane University School of Medicine
Nancy Gary, M.D.
  Albany Medical College
Beatrix Hamburg, M.D.
  Mt. Sinai Hospital
Murray M. Kappelman
  The University of Maryland School of Medicine
Ralph W. Muller
  University of Chicago Hospitals
Herbert Pardes, M.D.
  College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
  University of Alabama School of Medicine
W. Vickery Stoughton
  Toronto General Hospital
Martha Vaughan, M.D.
  National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
D. Kay Clawson, M.D.
  University of Kansas Medical Center School of Medicine

Chairman-Elect
David H. Cohen, Ph.D.
  Northwestern University

Immediate Past Chairman
John W. Colloton, M.A.
  University of Iowa Hospitals and Clinics

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
  George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Ernst R. Jaffé, M.D.
  Albert Einstein College of Medicine

Chairman, Council of Teaching Hospitals
Gary Gambuti
  St. Luke's-Roosevelt Hospital Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, N.W., Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.
Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

iV

INVITED ARTICLE

PETER A. WEIL, Ph.D.

Job Turnover of CEOs in Teaching and Nonteaching Hospitals

Abstract—Using data published annually by the American Hospital Association, this research compares the turnover rates of chief executive officers (CEOs) at approximately 800 teaching and approximately 4,700 nonteaching hospitals. (In this study, teaching hospitals were all hospitals with any reported affiliation with a medical school. About half of the hospitals studied had a major commitment to teaching, with several residency programs and a substantial number of residents in each program.) For each year studied (1981-1987), the teaching hospital CEOs experienced less turnover than did nonteaching hospital CEOs. While the turnover of CEOs from nonteaching hospitals showed a linear increase each year, culminating at 25.2% in 1987, the turnover of CEOs from teaching hospitals showed a more erratic pattern, resulting in 18.7% that year. A group of hospitals was identified that accounted for a large proportion of all recent CEO turnover. For example, 16% of the teaching hospitals experienced two or more CEO turnovers between 1981 and 1986. But these "higher-turnover" hospitals accounted for almost half of all teaching hospital turnovers. The higher-turnover hospitals were disproportionately found among non-federal government hospitals, in regions outside the Midwest, and in large cities; these hospitals also experienced higher costs per patient day and lower operating margins than did the lower-turnover hospitals. The research suggests that turnover may be accounted for, in part, by the comparative performance of a hospital relative to other hospitals of its type. Thus, whether hospitals experience stable, increasing, or declining operating margins, the higher-turnover hospitals typically underperform relative to other hospitals. *Acad. Med.* 65(1990):1-7.

Increasingly we hear about hospitals whose CEOs leave their posts—sometimes after many years of service—and we may wonder why there is so much turnover at the top. To study this problem in detail, three organizations[1] came together to begin a thorough study of CEO turnover. The first phase of the study, already completed, describes the extent of hospital CEO turnover over a six-year period, the kinds of hospitals experiencing higher numbers of turnovers, characteristics of the CEOs who left, and where they went.[1] The next phase of the study is currently under way; its objective is to develop a greater understanding of the reasons underlying CEO turnover.

CEO turnover is usually regarded as a negative phenomenon. It creates a short-term crisis in some hospitals because it interrupts the normal flow of activities, and many projects are put off until the new CEO is appointed. Not only do many hospitals spend a lot of time, effort, and money to find a suitable replacement, but other positions often will be vacated as the new CEO brings in a new team.

Of course, the departing CEO may feel angry, betrayed, or humiliated. Often the loss of the job means family displacement and financial pressure. Turnover often prevents CEOs from realizing their vision of where the hospital should be going, and it can short-circuit their careers.

But CEO turnover can bring positive outcomes for both the hospital and the CEO. For the hospital, it provides opportunities to think about new directions, and about developing new relationships with other providers, purchasers, government officials, and the community at large. It also provides the opportunity to abandon old unsuccessful strategies and bury hatchets from previous battles.

For many executives, too, turnover can mean a step up to a more challenging hospital CEO position, or it can lead to a shift to a non-hospital position with new and interesting challenges. For some it represents an opportunity to take a less demanding position; for others, it represents a time for career closure and a leisure-oriented lifestyle.[2]

A certain amount of CEO turnover is necessary and good. One of the opportunities this study provided was to identify hospital industry norms and to track these patterns over time. In this article, I focus on three questions.

Dr. Weil is the director of research and public policy at the American College of Healthcare Executives, Chicago, Illinois.

Correspondence should be addressed to Dr. Weil at the American College of Healthcare Executives, 840 North Lake Shore Drive, Chicago, IL 60611. Reprints will not be made available.

One: What are the recent CEO turnover rates at teaching and nonteaching hospitals?

Two: Is there such a thing as a "higher-turnover" hospital—a hospital that can't seem to hold onto its CEO for long?

Three: What are some of the distinguishing characteristics of CEOs who leave their positions?

I conclude by suggesting what part of the total turnover phenomenon I think is caused by a "ripple effect" as CEOs are recruited to other CEO slots, and speculate about some of the other possible causes of hospital CEO turnover.

Sources of Data

To answer the first question of the study, I used data from a six-year period, 1981–1987; to answer the second, I used data from a five-year period, 1981–1986. All these data were collected by the American Hospital Association and published in its annual *Guide to the Health Care Field*,[3] a comprehensive listing of the nation's hospitals. CEO turnover was denoted when the reported name of the administrator changed from one year to the next. This method of reporting understates turnover when there is more than one CEO change in a year.

To investigate the third question, I also used the biographical database of the American College of Healthcare Executives (hereafter referred to as "the College") for the years 1981 and 1987. This society is the largest professional association of healthcare executives, three-fourths of whom are hospital administrators; over half the hospital CEOs in the United States belong to the College.

I restricted the analysis regarding the first two questions to short-term, general medical and surgical, and non-federal-government hospitals. In these categories, the teaching hospitals were defined as hospitals having any affiliation with a medical school. There were 801 such hospitals in 1986–87. (About half of these had a major commitment to teaching, with several residency programs and a substantial number of residents in each program.) During the same year, there were 4,724 nonteaching hospitals in the categories previously mentioned.

However, slightly smaller numbers of both teaching and nonteaching hospitals are reported when I describe the higher- and lower-turnover hospitals in the Results section of this article. This is because it was necessary to study trends in hospital CEO turnover over a period of years (1981–1986) in order to examine the second question (whether there is a "higher-turnover" hospital); hospitals that failed to report the name of the CEO or that were not listed in each of the years studied were excluded from the analysis.

I examined the following hospital characteristics to try to distinguish the higher-turnover hospitals from the lower-turnover ones: bedsize, membership in a hospital system, type of ownership, region in which the hospital was located, size of the community where the hospital was located, level of admissions, occupancy rate, cost per patient day, and operating margin.

Prediction Difficulties

For a variety of reasons, it is difficult to predict the turnover rates of CEOs at teaching and nonteaching hospitals. On the one hand, several factors argue for teaching hospitals to experience higher CEO turnover rates than their nonteaching counterparts. First, most teaching hospitals serve a high proportion of poor and uninsured patients.[4] This would leave them open to financial losses and consequent dissatisfaction with management. Second, some teaching hospitals have full-time medical staffs—people who are powerful, entrenched, and tenured. Obviously, when crossed—for whatever reason—their clout might very well prevail over that of the executive, who may have an employment contract but is not tenured. Third, CEOs in around half the teaching hospitals in this study's sample are pledged to pursue a complicated mission: teaching, research, and patient care. Juggling these three key objectives while keeping the hospital financially solvent may require superhuman abilities.

But several other features of teaching hospitals argue for their CEOs to enjoy less turnover than their counterparts in nonteaching hospitals. First, unlike nonteaching hospitals, Medicare has allowed them certain cost adjustments —notably a 7.5% allocation for teaching costs, for the inefficiencies necessarily introduced in the teaching process, and for the severity of illnesses experienced by patients there. So when nonteaching hospitals are totally at the mercy of prospective payment, teaching hospitals have a little compensation for their higher costs. Second, because many medical staffs in major teaching hospitals are employed full-time, they have ample opportunity to interact with executives, are themselves aware of the difficulty of administering departments and the various services, and are strongly invested in the hospital's well-being. Ultimately, these physicians may feel less alienated from hospital executives and may indeed support them to a greater extent than would physicians in nonteaching settings.

Third, the very complexity of the teaching hospital's mission makes the evaluation of managers difficult. Thus, by dint of the near-impossibility of fulfilling all three objectives of teaching, research, and patient care, managers might be given an overall favorable evaluation even if they may have performed below expectations in some areas. Fourth, teaching hospitals often are larger; their operating margins tend to be higher than those of nonteaching hospitals.[5] Analogous studies of school superintendents showed that those in larger districts were less likely to experience turnover than were those in small districts.[6] Finally, boards of trustees of larger hospitals are more likely to use predetermined criteria for selecting as well as evaluating the CEO; thus there might be fewer board-initiated terminations.[7]

When I considered the kinds of individuals who are the CEOs of teaching hospitals compared with the CEOs in nonteaching hospitals, my experience led me to predict that the teaching hospital CEOs would be at the pinnacle of their careers, would have more experience in health administration, and would be older. Consequently, their leaving might be for the purpose of retirement. On the other hand, nonteaching hospital CEOs might be on their way to greater challenges and thus more likely to move on to a different hospital. Thus, because the two types of hospitals could represent different stages in the careers of CEOs, it is impossible to predict what impact individual differences of CEOs might have on overall turnover rates.

Findings

The findings concerning the three questions of this study were highly illuminating and suggest that turnover may be as much a product of the general hospital context in which CEOs are located as it is a product of CEOs' personal career plans.

One: CEO Turnover Rates

Figure 1 shows that the 749 teaching hospitals for which adequate data were available experienced lower rates of CEO turnover every year for the six years of the study than did the 4,539 nonteaching hospitals for which there were data. Also, an erratic pattern is evident for teaching hospital CEOs. After a significant increase between 1981 and 1982, the turnover rate actually decreased between 1983 and 1985, only to rise sharply in the most recent time period. However, the year-to-year trends within the teaching hospitals are not as statistically stable as is true for the nonteaching group, since they are based on data from a smaller group of hospitals.

For both the teaching and the nonteaching hospitals, the biggest single-year increase in turnover was during the 12-month period immediately before the introduction of the prospective payment system (that is, from 1981–82 through 1982–83). The average annual rate of increase for the last five years studied was 2.3% for the teaching hospitals; for the nonteaching hospitals, it was 4.9%.

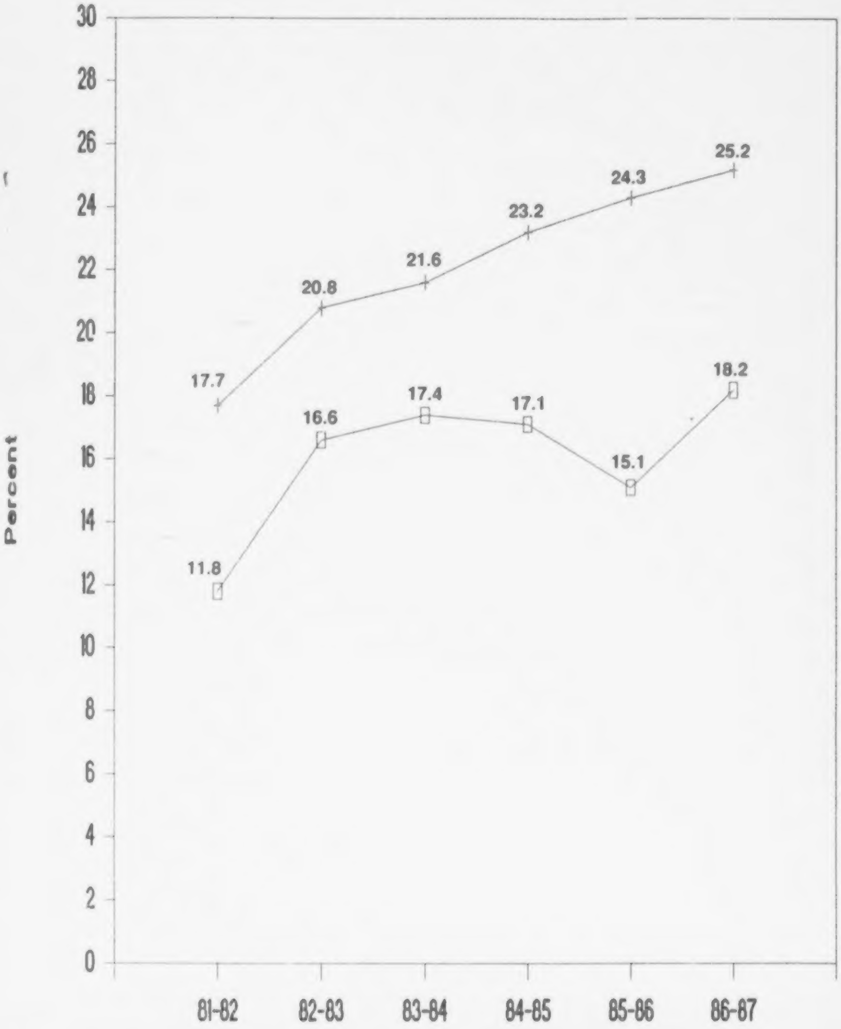


Figure 1. **Job turnover of chief executive officers in 749 U.S. teaching hospitals and 4,539 nonteaching hospitals, 1981–1987. Teaching hospitals were defined as those having any degree of affiliation with a medical school. For both the teaching and the nonteaching hospitals studied, the biggest single-year increase in turnover was during the 12-month period from 1981–82 through 1982–83, which was the time immediately before the introduction of the prospective payment system. Average annual rates of increase for the last five years studied were 2.3% for the teaching hospitals and 4.9% for the nonteaching hospitals. *Rectangles* represent teaching hospitals; *crosses* represent nonteaching hospitals.**

Two: "Higher-Turnover" Hospitals

One advantage of looking at turnover for a period of years is identifying hospitals with higher rates of turnover. Using the data from the American Hospital Association, I examined hospital turnover for each year between 1981 and 1986. (As already noted, the hospitals that failed to report the name of the CEO or that were not listed in each of the five years were excluded from the analysis concerning the second question of this study.) Hospitals were classified into three groups: those with no turnover in the five-year period, those with one turnover, and those with two or more turnovers. Hospitals with no turnover or with one turnover were designated "lower-turnover" hospitals, while those experiencing two or more turnovers were labeled "higher-turnover" hospitals.

1. A small percentage of hospitals accounted for most hospital CEO turnover. In the five-year period beginning in 1981–82 and ending in 1985–86, approximately 46% of all teaching hospitals had no CEO turnover, while 38% of them experienced one change in the CEO position. What is critical is that 16% of the hospitals that experienced two or more turnovers—the higher-turnover teaching hospitals—accounted for 49% of all turnovers.

Among the nonteaching hospitals, a similar pattern prevailed. Between 1981–82 and 1986–87, 37% experienced no CEO turnovers, 36% had one turnover, and 27%—the higher-turnover nonteaching hospitals—had two or more turnovers. This last group—the higher-turnover nonteaching hospitals—accounted for 64% of all turnovers.

While both a percentage of both the teaching and the nonteaching hospitals were composed of higher-turnover hospitals, a notable difference between the two groups was that the higher-turnover hospitals were a smaller proportion of the teaching group than they were of the nonteaching group. The implication: getting a CEO job in a teaching hospital means that the CEO is less likely to change jobs. The chances of a CEO's leaving a position in a higher-turnover teaching hospitals were one in six during the years studied, but in nonteaching hospitals, these same chances were about one in four.

2. CEOs were much more likely to experience turnover in the higher-turnover hospitals. What did serving in higher- or lower-turnover hospitals mean to the careers of hospital CEOs? There was a 50% probability that CEOs at the helms of higher-turnover hospitals (whether teaching or nonteaching) would leave their organizations in the following year, making their average tenure only two years.

In contrast, executives who headed lower-turnover teaching hospitals had only a 9% probability of leaving in the following year, giving them an average tenure of 11 years. CEOs in lower-turnover nonteaching hospitals had a 10% probability of leaving in the following year, giving them an average tenure of ten years.

3. All nine hospital characteristics used in this study helped distinguish the higher-turnover nonteaching hospitals from those with lower turnover. Listed in the following paragraphs are the nine hospital characteristics that served to identify where the CEO turnover rates in nonteaching hospitals were higher. The nine variables were determined to be statistically significant ($p < .05$) using chi-square tests for categorical variables and analysis of variance for continuous variables.

SMALLER HOSPITALS. From 1981 to 1986, smaller hospitals experienced higher turnover. In 1986, for example, the nonteaching hospitals with no turnover averaged 137 beds. The hospitals with one turnover had an average of 130 beds, while those with two or more turnovers had an average size of 109 beds. Stated another way, 73% of the hospitals with fewer than 50 beds had two or more turnovers, compared with 19% of hospitals with over 400 beds.

SYSTEM HOSPITALS. Thirty-six percent of the nonteaching hospitals that were part of a system had higher turnover. In contrast, only 19% of the nonteaching hospitals that were freestanding experienced higher turnover, a factor that may be due in part to intrasystem promotions or reassignments.

INVESTOR-OWNED HOSPITALS. Between 1981 and 1986, 49% of the nonteaching investor-owned hospitals experienced two or more CEO turnovers. This compares with 22% of the not-for-profit church hospitals, 21% of the not-for-profit secular hospitals, and 27% of the state and local government-owned hospitals that were higher-turnover hospitals.

HOSPITALS IN THE SOUTH AND WEST. Higher turnover among nonteaching hospitals was concentrated in the West—37% of the hospitals had two or more turnovers in the five-year period—and in the South, where 31% of the hospitals were higher-turnover ones.

HOSPITALS IN THE LARGEST CITIES. The hospitals located in cities above 156,000 population experienced higher turnover. Thirty-six percent of the hospitals in the 100 largest cities had two or more CEO turnovers. In contrast, small cities had the lowest percentage of higher turnover—averaging 22%. About 27% of the higher-turnover hospitals were in nonmetropolitan and suburban sites.

HOSPITALS WITH LOWER ADMISSIONS. Between 1981 and 1986, the nonteaching hospitals with lower admissions experienced higher turnover. In 1986, for example, the hospitals with no turnover averaged 4,566 admissions, and the hospitals with one turnover experienced 4,132 admissions. The hospitals with more turnovers averaged only 3,188 admissions.

HOSPITALS WITH LOWER OCCUPANCY RATES. The nonteaching hospitals with higher turnovers had occupancy rates 7% lower than the occupancy rates of the hospitals with no turnover between 1981 and 1986. In 1986, for example, the hospitals with no turnover had an average occupancy rate of 55%; those with one turnover averaged 52%. But higher-turnover hospitals averaged 48% that year.

HOSPITALS WITH HIGHER COSTS. Again, a consistent pattern was seen for the higher-turnover hospitals to experience higher costs than those experienced by other, lower-turnover hospitals. For example, in 1986 the average total cost per patient day for the no-turnover hospitals was $410; for the one-turnover hospitals it was $436. But the higher-turnover hospitals averaged $467 per day.

HOSPITALS WITH LOWER OPERATING MARGINS. Using unpublished data from the American Hospital Association, Figure 2 shows the operating margins for nonteaching hospitals by CEO turnover status. For each year between 1982 and 1987, a gradient appears. The highest operating margins are in the zero-turnover hospitals, the next-highest margins are in the one-turnover hospitals, and the lowest margins are in the hospitals that experienced two or more turnovers between 1981 and 1986.

4. Only five hospital characteristics used in this study helped distinguish the higher-turnover teaching hospitals from those with lower turnover. The useful characteristics were ownership, region, size of community, costs, and operating margin. Listed in the following paragraphs are the five hospital characteristics that served to identify where CEO turnover rates in teaching hospitals were higher. These variables were determined to be statistically significant by using the same statistical procedures previously described for the nonteaching hospitals.

NON-FEDERAL GOVERNMENT HOSPITALS. Between 1981 and 1986, 24% of the non-federal government hospitals experienced higher turnover, compared with 11% of the not-for-profit church hospitals and 17% of the not-for-profit secular hospitals.

HOSPITALS OUTSIDE THE MIDWEST. The higher-turnover teaching hospitals tended to lie outside the midwestern region, where only 11% had multiple CEO turnovers between 1981 and 1986. In contrast, the higher-turnover hospitals constituted 18% of hospitals in the Northeast, 18% of the hospitals in the West, and 20% in the South.

HOSPITALS IN LARGE CITIES. The higher-turnover hospitals were concentrated in the largest cities (over 156,000 population), where they constituted 21% of all hospitals studied. They represented only 11% of the hospitals in smaller cities, 14% of those in suburban areas, and 15% of the hospitals in nonmetropolitan areas.

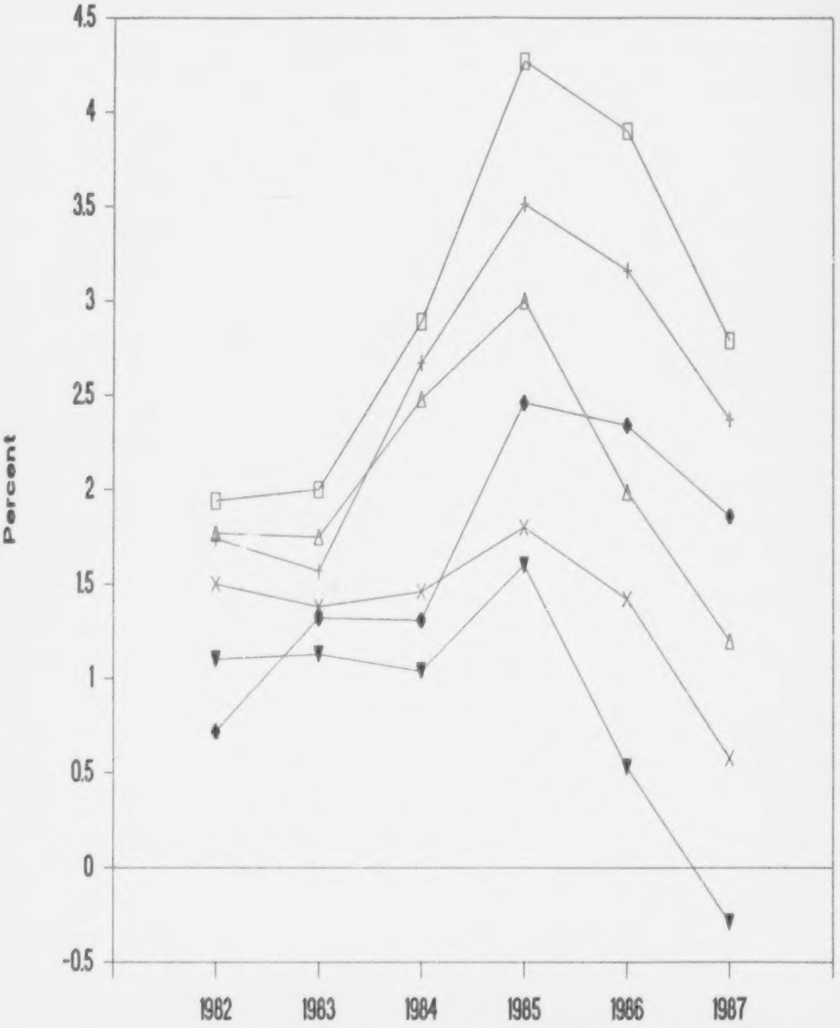


Figure 2. **A comparison of the median operating margins and job turnovers of chief executive officers (CEOs) in 749 U.S. teaching hospitals and 4,539 nonteaching hospitals, 1981 – 1986, and 729 teaching and 4,452 nonteaching hospitals in 1987. Teaching hospitals were defined as those having any degree of affiliation with a medical school. After 1985, both groups of hospitals experienced declining operating margins; however, the teaching hospitals with two or more CEO turnovers experienced a lower rate of decline in operating margins than did any other type of hospital. Also, consistent with their lower turnover rates, the teaching hospitals had higher operating margins than did the nonteaching hospitals. Hollow symbols indicate no CEO turnover in teaching hospitals (*rectangles*) and nonteaching hospitals (*triangles*). Line symbols indicate one CEO turnover in teaching hospitals (*crosses*) and nonteaching hospitals (*Xs*). *Solid symbols* indicate two or more CEO turnovers in teaching hospitals (*diamonds*) and nonteaching hospitals (*inverted triangles*).**

HOSPITALS WITH HIGHER COSTS. From 1981 to 1986 the hospitals with higher costs per patient day experienced higher turnover. For the hospitals with no CEO turnover during the same time period, the average total cost per patient day was $546 in 1986; for those with one turnover, the average total cost was $570. But teaching hospitals with two or more turnovers produced the highest average cost: $587.

HOSPITALS WITH LOWER OPERATING MARGINS. As was true for the nonteaching hospitals, a gradient in operating margin appeared for each year (1982–1987) between the hospitals with zero, one, and two or more CEO turnovers. The highest margins were seen in zero-turnover hospitals; the lowest margins were in the two-or-more-turnover hospitals. After 1985, when operating margins peaked, each type of hospital experienced declining operating margins. However, the teaching hospitals with two or more CEO turnovers experienced a lower rate of decline than did any other type of hospital. (See Figure 2.) Also, consistent with their lower turnover rates, the teaching hospitals had higher operating margins than did the nonteaching hospitals. Tests of differences between means showed significantly lower operating margins between the two-or-more-turnover hospitals and others in all cases except for teaching hospitals in 1983 and 1986.

To control for the fact that many of these variables were interrelated, I conducted an ordinary least-squares regression analysis. Variables that were statistically (at $p<.10$ or less) significant in predicting being a higher-turnover teaching hospital were metropolitan location, membership in a multihospital system, and low occupancy rate. Also, being under not-for-profit church auspices is negatively related to being a higher-turnover teaching hospital. The volatility of the mean operating margin probably accounts for its insignificance as a predictor. Instead, low occupancy rate is significantly associated with teaching hospitals that experience high CEO turnover.

Three: CEO Characteristics

Even though certain hospitals seem to experience CEO turnover more than others, turnover is experienced by an individual as well. To help identify whether specific types of individuals are more susceptible to turnover, the existing database of the American College of Healthcare Executives ("the College") was used. That database includes a comprehensive biography of approximately half of the nation's hospital CEOs.

The College is disproportionately composed of CEOs of larger (over 200 beds), not-for-profit, general medical, surgical, and children's hospitals, as well as hospitals that are system members. In addition, these executives are more heavily represented in the nation's teaching hospitals, metropolitan areas, and the New England, Middle Atlantic, and East North Central regions.

Using the College's database for the years 1981 and 1987, it became clear that turnover of hospital CEOs in 1987 was different than it had been as recently as 1981. In 1981, for example, turnover was higher among younger executives, people under age 35 who had been in their positions for less than the median period of CEO service (4.25 years, a relatively short period). But today, turnover is being experienced by executives in all age groups and, in fact, is slightly higher among CEOs who have been in their positions longer than average. (College-affiliated CEOs have about the same tenure today as they did in 1981. Then, the mean tenure for departing CEOs was 7.1 years; in 1987, it was 6.5 years.)

The data indicate that the CEOs who left their positions in 1987 early (that is, before the median period of 4.25 years) not only could be any age, but had the following characteristics: (1) they more often than not were unmarried, (2) they were not involved in preceptorships—that is, taking on residents in health administration, (3) they had short tenures in their previous positions, and (4) they had served as CEOs in locations other than their birthplaces.

On the other hand, it was difficult to characterize the CEOs who left after relatively long (more than 4.25 years) tenures. The data indicated no important differences between them and those who stayed on, including demographic, career, and hospital characteristics.

A real advantage of studying this subgroup of hospital CEOs from the College data is to obtain some estimate of how much turnover occurs within their own ranks. In 1987, 40% of the College CEOs who left their positions as heads of hospitals moved to become CEOs of other hospitals (30.5%) or CEOs of hospital systems (9.5%). This gives a new slant to the turnover statistics presented earlier, since it tells us that turnover in hospitals creates a ripple effect. For instance, in 1987 academic hospitals experienced an 18.2% turnover rate. But, because of the desirability of these hospitals, it is likely that many of the positions were filled by existing CEOs elsewhere in the hospital industry. Thus, the 18.2% can be reasonably partitioned into a "natural" category—turnover due to retirements, force-outs, etc., which is 11%—and a "ripple effect" category of 7.3%, which is due to the recruitment of other hospital CEOs. An 11% natural turnover rate is not an outrageous one; it corresponds to a tenure of over nine years.

Summary and Discussion

The research reported here lends some insight into the kinds of hospitals that experience high rates of CEO turnover. Turnover in teaching hospitals is highest in non-federal government hospitals, hospitals outside the Midwest, and hospitals in large cities. Higher costs per patient day and low operating margins also characterize the higher-turnover teaching hospitals. The inconsistent relationship between net operating margin and CEO turnover rates in both teaching and nonteaching hospitals suggests that turnover represents factors other than the "bottom line," a finding noted in previous research outside the healthcare sector.[8–10] Nevertheless, the higher-turnover hospitals typically underperformed in profitability relative to the other hospitals studied. Thus, comparison of a hospital's performance with those of others of its type may trigger the CEO's departure.

Of course this secondary analysis of

existing data did not permit us to investigate other potentially more sensitive variables related to CEO turnover. For example, some hospitals have outdated governance and administrative structures that may underlie poor operational performance of these organizations. These hospitals may serve as a "proving ground" to determine which CEOs will be selected by more financially stable institutions. Likewise, the level of executive pay, while related to hospital size, probably provides some independent contribution to explaining CEO turnover.

Despite a comprehensive biographical database on the individual careers of the College-affiliated hospital CEOs, few features distinguished the leavers from the stayers. The most recent data show that the CEOs who are more likely to leave their positions after having served less than the average length of time are unmarried, lack involvement as preceptors to health administration students, have histories of short tenures in previous positions, and reside in states away from their birthplaces. In contrast to historical patterns, these CEOs are not younger than those who do not leave their positions early; they can be of any age group.

While these five variables capture significant differences between the leavers and the stayers, no significant relationships were found to characterize the longer-tenured leavers compared with the stayers. And it is the longer-staying College CEOs who currently experience a higher rate of turnover—17.6% compared with those in their positions for a shorter-than-average time—15.1%.

More precise hypotheses will be studied in the next phase of this research. For example, we plan to examine a sample of CEOs to study their values, working relationships, management styles, and expectations to remain or leave the hospital. Then, a year later, the study will determine who have left their positions, where they went, and why. We expect the observations from such research will allow a better understanding of the causes of turnover. Realistically speaking, however, we know that hospital CEO turnover will always be partly inscrutable. Difficult-to-measure factors such as personal values, organizational culture, ambition, and charisma influence the tenures of hospital CEOs. Moreover, this next phase of the research will not answer the question of whether or not the turnover that occurred brought about positive outcomes for the individual hospital or its community. Additional longitudinal research may help to answer such key questions.

The author thanks C. Thomas Smith for his comments on an earlier version of the manuscript, and John Timmerberg and Lloyd Wackerling for assisting in computer programming.

References

1. American College of Healthcare Executives, American Hospital Association, and Heidrick and Struggles, Inc. *Hospital CEO Turnover: National Trends and Variations (1981-1987).* Chicago, Illinois-ACHE, AHA, and H&S, Inc., 1989.
2. Dalton, Dan, and Todor, William R. Turnover Turned Over—An Expanded and Positive Perspective. *Acad. Manage. Rev.* **4**(1979):225-235.
3. American Hospital Association. *Guide to the Health Care Field.* 1988 ed. Chicago, Illinois: American Hospital Association, 1988.
4. Sloan, Frank A., Balbona, J., and Mullner, Ross. Identifying the Issues: A Statistical Profile. In *Uncompensated Hospital Care,* F. A. Sloan, J. F. Blumstein, and J. M. Perrin, eds., pp. 16-53. Baltimore, Maryland: Johns Hopkins University Press, 1986.
5. Cleverley, William O. *Hospital Industry Financial Report 1983-1987.* Westchester, Illinois: Healthcare Financial Management Association, 1988.
6. March, James C., and March, James G. Almost Random Careers: The Wisconsin School Superintendency, 1940-1972. *Admin. Sci. Q.* **22**(1977):377-409.
7. Alexander, Jeffrey, and Brooks, Deal C. New Dimensions in Board-CEO Relations. *Trustee* **39**(June 1986):24-27.
8. Allen, Michael P., Panian, Sharon K., and Lotz, Roy C. Managerial Succession and Organizational Performance: A Recalcitrant Problem Revisited. *Admin. Sci. Q.* **24**(1979):167-180.
9. Salancik, Gerald R., and Pfeffer, Jeffrey. Effects of Ownership and Performance on Executive Tenure in U.S. Corporations. *Acad. Manage. J.* **23**(1980):653-664.
10. Wagner, W. Gary, Pfeffer, Jeffrey, and O'Reilly, Charles A., III. Organizational Demography and Turnover in Top Management Groups. *Admin. Sci. Q.* **29**(1984):74-92.

A3273-8

FEBRUARY 1990

ACADEMIC MEDICINE

Rising Cost of NIH-Funded Research

Calls for Change in Academic Medicine

Three Easy Pieces

A Focus for Change

Physician Education

AAMC Annual Awards

"Certifimania" in Medical Education

Teaching Critical Thinking

Women's and Men's Research Activities

Information-Gathering Strategies

Improving Oncology Education

Role of Nonphysician Consultants

VOLUME 65 • NUMBER TWO

FEBRUARY 1990
VOLUME 65
NUMBER TWO

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Invited Articles

The Rising Cost of NIH-Funded Biomedical Research? *by Thomas J. Kennedy, Jr.*, M.D. 63

Three Easy Pieces *by Robert* G. *Petersdorf*, M.D. 73

Graduate Medical Education: Focus for Change *by Spencer Foreman*, M.D. 77

The Education of the Physician *by* D. *Kay Clawson*, M.D. 84

Ideas for Medical Education

Teaching Critical Thinking in the Context of Substance Abuse in a Psychiatry Clerkship *by Martha* G. *Regan-Smith*, M.D., *and Donald* A. *West*, M.D. 89

Awards

Annual Awards of the Association of American Medical Colleges 90

Commentary

20 Years of the VA's Distinguished Physician Program *by James* E. *Eckenhoff*, M.D. 94

From the President 96

AAMC Focus 97

National Policy Perspectives 98

Letters to the Editor

Health Care and the Private Sector in the Third World *by Raymond* A. *Smego, Jr.*, M.D., M.P.H., *and* *Peter* V. *Barrett*, M.D. 100

The Hippocratic Oath and Medical Students *by* H. *Steven Moffic*, M.D., *Timothy Bayer*, M.D., *and* *John Coverdale*, M.D. 100

English Proficiency in Multiple-Choice Questions *by* Dr. *Asoka* S. *Dissanayake*, Dr. *Basil* A. *Ali*, *and* Dr. *Usha Nayar* 101

Research Reports

ARTICLES

Comparing Research Activities of Women and Men Faculty in Departments of Internal Medicine *by Barbara* A. *Levey*, M.D., *Nancy* O. *Gentile*, H. *Paul Jolly*, *Ph.D.*, *Harry* N. *Beaty*, M.D., *and Gerald* S. *Levey*, M.D. 102

Comparing Information-Gathering Strategies of Medical Students and Physicians in Diagnosing Simulated Medical Cases *by Nancy* S. *Palchik*, *Ph.D.*, *Fredric* M. *Wolf*, *Ph.D.*, *James* T. *Cassidy*, M.D., *Robert* W. *Ike*, M.D., *and Wayne* K. *Davis*, *Ph.D.* 107

Improvement of Oncology Education at the University of Washington School of Medicine, 1984–1988 *by* W. *Archie Bleyer*, M.D., *Daniel* D. *Hunt*, M.D., *Jan* D. *Carline*, *Ph.D.*, *Michael Trevisan*, *M.Ed.*, *Suzanne Johnson*, *and Theodore* J. *Phillips*, M.D. 114

The Role of Nonphysician Consultants as Health-Care Educators in Postgraduate Programs of Anesthesiology *by Henry Rosenberg*, M.D., *and Binnie Polonsky*, M.P.H. 119

How Medical School Faculty Regard Educational Research and Make Pedagogical Decisions *by Marc* S. *Nelson*, M.D., *Barbara* L. *Clayton*, M.A., *and Roberto Moreno*, M.A. 122

SHORT PAPER

Informed Consent and the "Medical Student Psychiatrist" *by Daniel* L. *Cohen*, M.D., *Rosslyn* W. I. *Kessel*, *Ph.D.*, *Laurence* B. *McCullough*, *Ph.D.*, *Aristide* Y. *Apostolides*, *Ph.D.*, *and Kelly* J. *Heiderich* 127

Bibliography 129

Publishing Services: **Hanley & Belfus, Inc.**
210 South 13th Street, Philadelphia, PA 19107

FEBRUARY 1990
VOLUME 65
NUMBER TWO

ACADEMIC MEDICINE

JOURNAL OF THE ASSOCIATION OF AMERICAN MEDICAL COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, M.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
Arizona Health Sciences Center

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
Georgetown University Hospital
Philip Anderson, M.D.
University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
Harvard Medical School
Eta Berner, Ed.D.
University of Alabama School of Medicine
Ruth Bulger, Ph.D.
National Academy of Sciences
Vincent A. Fulginiti, M.D.
Tulane University School of Medicine
Nancy Gary, M.D.
Albany Medical College
Beatrix Hamburg, M.D.
Mt. Sinai Hospital
Murray M. Kappelman, M.D.
The University of Maryland School of Medicine
Ralph W. Muller
University of Chicago Hospitals
Herbert Pardes, M.D.
College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
University of Alabama School of Medicine
W. Vickery Stoughton
Toronto General Hospital
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
David H. Cohen, Ph.D.
Northwestern University

Chairman-Elect
William T. Butler, M.D.
Baylor College of Medicine

Immediate Past Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center School of Medicine

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Joe Dan Coulter, Ph.D.
University of Iowa College of Medicine

Chairman, Council of Teaching Hospitals
Raymond G. Schultze, M.D.
University of California, Los Angeles Medical Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.
Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

THOMAS J. KENNEDY, JR., M.D.

The Rising Cost of NIH-Funded Biomedical Research?

Abstract—During the last decade, total appropriations for the NIH have grown in current as well as constant dollars. Constant dollar expenditures for indirect costs and research project grants have increased, as also has the number of the latter, while such expenditures for research centers, training, and research contracts have shrunk. The most impressive redistribution in emphasis has been toward traditional research project grants (R01s). The size of the average R01 award, discounted for inflation, has grown at an annual rate of 1.1% during the last decade and 1.3% since fiscal year (FY) 1970; that of the average research program project (P01) has declined over the same periods, after a slight rise in the early 1970s. Factors contributing to the modest rise in the real (constant-dollar) size of the average R01 are explored. The regularity with which current-services-requirements estimates for the NIH exceed inflation reflects real growth in the program, particularly in the category of research project grants; the artifact of basing calculations on the post- rather than pre-"negotiated" levels of awards in the "current" year; and the extent to which the project periods of awards have been extended. The effect of lengthening project periods is slow to become manifest, but inexorably swells the pool of non-competing awards; decisions in this area undertaken in 1985, and continued at least through FY 1988, could very significantly increase current services requirements in FYs 1991 and 1992. *Acad. Med.* 65(1990):63–73.

For almost a decade, the federal government has been operating under a fiscal policy that has resulted in novel restraints and limitations. An examination of what effects that policy has had on the federal support of health research and what the prospects might be for future support seems timely. Since the bulk of federal funding for health research is channeled through the National Institutes of Health (NIH), the analysis reported in the present article has focused on the programs of that agency and has relied heavily on NIH publications as well as direct communication with NIH staff. The interpretations and speculations are entirely personal.

Understanding the fiscal climate that has prevailed in the nation during recent years is essential for evaluating historic and prospective changes in the patterns of federal support for biomedical research programs. A useful place to begin is with the major reduction in tax rates that occurred in 1982, thereby reducing revenues relative to outlays, with a consequent increase in the deficit.[1] (See Table 1.) This is most clearly evident when revenues and outlays are expressed as percentages of a rising Gross National Product (GNP) but is also seen in dollar terms where growth in outlays outpaced that for revenues by a substantial margin.

The absolute value of the deficit had reached such alarming proportions that Congress moved to control it through The Balanced Budget and Emergency Deficit Control Act of 1985, better known as the Gramm-Rudman-Hollings legislation. The essential mandate of this law is systematic reduction of the annual deficit, scheduled to reach zero in fiscal year (FY) 1992. It does not specify how the debt should be reduced but, without any parallel move to increase revenues, the only device to achieve the goal has been to curtail expenditures. Severe pain was postponed for a while, but the vise has steadily tightened. Many federal programs—Preventive Health Services block grants, the National Center for Health Services Research, the National Health Service Corps, and public housing—have experienced real and deep cuts. Others that may have enjoyed modest increases—the Alcohol, Drug Abuse, and Mental Health Administration, the Maternal and Child Health and Primary Care block grants—have failed, by wide margins, to keep up with inflation. The commitments of entitlement programs have been eroded, for Medicare by perhaps as much as $40 billion, thereby limiting benefits even as the pool of those who are eligible grows.

Most members of Congress know that the fat is long gone and sinew and muscle are now being sacrificed. But successive administrations remain intransigent on raising revenue and Congress has fallen in line. Whether the politicians are leading the public or vice versa, the current

---

Dr. Kennedy is an associate vice president, Association of American Medical Colleges.

Correspondence and requests for reprints should be addressed to Dr. Kennedy, AAMC, 1 Dupont Circle, N.W., Washington, DC 20036.

Table 1

The Budget of the United States: Revenues, Outlays, and Deficits, Fiscal Years 1962–1988*

Fiscal Year	As Percentage of GNP			In Billions of Dollars		
	Revenues	Outlays	Deficit or Surplus	Revenues	Outlays	Deficit or Surplus
1962	17.9	19.2	−1.3	99.7	106.8	−7.1
1963	18.1	18.9	−0.8	106.6	111.3	−4.8
1964	17.9	18.8	−0.9	112.6	118.5	−5.9
1965	17.4	17.6	−0.2	116.8	118.2	−1.4
1966	17.7	18.2	−0.5	130.8	134.5	−3.7
1967	18.7	19.8	−1.1	148.8	157.5	−8.6
1968	18.0	21.0	−3.0	153.0	178.1	−25.2
1969	20.1	19.8	0.3	186.9	183.6	3.2
1970	19.5	19.8	−0.3	192.8	195.6	−2.8
1971	17.7	19.9	−2.2	187.1	210.2	−23.0
1972	18.0	20.0	−2.0	207.3	230.7	−23.4
1973	18.0	19.2	−1.2	230.8	245.7	−14.9
1974	18.6	19.0	−0.4	263.2	269.4	−6.1
1975	18.3	21.8	−3.5	279.1	332.3	−53.2
1976	17.6	21.9	−4.3	298.1	371.8	−73.7
1977	18.4	21.2	−2.8	355.6	409.2	−53.6
1978	18.4	21.1	−2.7	399.6	458.7	−59.2
1979	18.9	20.6	−1.6	463.3	503.5	−40.2
1980	19.4	22.1	−2.8	517.1	590.9	−73.8
1981	20.1	22.7	−2.6	599.3	678.2	−78.9
1982	19.7	23.8	−4.1	617.8	745.7	−127.9
1983	18.1	24.3	−6.3	600.6	808.3	−207.8
1984	18.1	23.1	−5.0	666.5	851.8	−185.3
1985	18.6	23.9	−5.4	734.1	946.3	−212.3
1986	18.4	23.7	−5.3	769.1	990.3	−221.2
1987	19.3	22.6	−3.4	854.1	1,003.8	−149.7
1988	19.0	22.3	−3.2	909.0	1,064.1	−155.1

*These data indicate the sharp increase in the deficit that occurred when taxes were reduced in 1982. (Source: Congressional Budget Office.)

policy—no new taxes—appears to have broad public support. Thus, for government programs, their beneficiaries and their advocates, the struggle to stay whole has been replaced by the struggle to stay alive.

In this circumstance, it is encouraging that appropriations for the NIH during the past decade have grown steadily in current dollars and also, except for a slump between FYs 1979 and 1982, in constant dollars.[2] However, in the pressure cooker that the annual appropriations process has become, generosity toward the NIH comes increasingly at the expense of other programs. The beneficiaries of, and advocates for, these programs have become both strident in protesting reductions and insistent on a larger share of the available funds. In this atmosphere the NIH costs have come under intense scrutiny.

The most recent challenge occurred in the spring of 1988 during the presentation of the president's budget for FY 1989. The director of the NIH stated that the agency would need a 16.1% increase (later corrected to 14.1%) in order simply to maintain the same level of program activity that was being supported in FY 1988. This aroused House Appropriations Committee interest in why the requirement was so much higher than that expected from inflation, and revived old misgivings about the seemingly inexorable increases in the average "cost" of a grant. When the dust had settled, the Department of Health and Human Services (DHHS) Inspector General was asked by the House Appropriations Committee to "review a sample of extramural grants to determine whether these costs are being well managed. To the extent possible, the Committee would request that this review include information on manpower utilization, salary growth and procurement practices."[3]

The rising cost of biomedical research has suddenly become a central public issue and could join the rising cost of medical care as a perennial one.

Definition of the Cost Problem

The precise nature of the dis-ease that prevails in Congress and elsewhere about research costs is still inchoate. Under the general rubric of cost, at least three discrete issues can be discerned: the average cost of a research grant, the total cost of the

NIH program, and the annual escalation in the current services requirements in excess of inflation.

Average Cost of a Grant

The problem of the average cost of a grant, as perceived by Congress, is that the unit cost for a research project grant (RPG) rose from $87,720 in FY 1978 to $184,431 in FY 1988.[4][*] Granting that "cost" is something of a misnomer, and that what is actually at issue is the amount that the NIH agreed to pay for the proposed scope of work, the increases are certainly impressive.

Magnitude of Increases

Averages can mask important details; in the interest of greater precision, more specificity is warranted. The major formal categories of the NIH budget[2] include research grants, research training grants, research and development contracts, and intramural research, as well as several others less germane to the topic under examination. The category of research grants comprises three subcategories: research project grants; research center grants, including General Clinical Research Centers and Biotechnology Research Centers; and other research grants, including career awards, Biomedical Research Support Grants (BRSGs), and Minority Biomedical Research Support Grants (MBRSGs).

The subcategory of research project grants — the average dollar size of these was just cited — is itself not homogeneous but includes a substantial number of discrete instruments of support: Traditional Research Project (R01) Awards; Research Program Project (P01) Awards; US-Japan Cooperative Medical Science Program (R22) Awards; New Investigator Research (R23) Awards, currently being phased out and replaced by FIRST (First Independent Research Support and Transition) (R29) Awards; Phase I and Phase II SBIR (Small Business Innovative Research) (R43 and R44) Awards; Research Project Cooperative (U01) Agreements; Outstanding Investigator (R35) Awards; MERIT (Method for Extending Research in Time) (R37) Awards; and Hazardous Substance Basic Research (P42) Awards.

Some of these types of awards are quite old, but many are of recent vintage. Their average "costs" in FY 1988 varied over a wide range — from under $50,000 for R23 and R43 awards to almost $1.0 million for P01s and P42s.[2,4] The R01s and P01s accounted for about 84% of all research project grant expenditures in FY 1988. The average award size for R01s increased over the decade FYs 1979-1988 from $78,117 to $155,415, a slightly more rapid rate than for the P01s, which grew from $570,377 to $850,950.[2,4]

Effect of Inflation

While these increases are large, they coincide with an epoch during which inflation was one of the country's major economic problems. Average prices can be discounted for inflation by the use of several indices: the implicit Gross National Product (GNP) deflator or some specific component thereof[5]; the Consumer Price Index (CPI)[6]; or the Biomedical Research and Development Price Index (BRDPI). This latter was developed for the NIH by the Bureau of Economic Analysis of Department of Commerce[7,8]; its validity has withstood challenge well, and accordingly it is used throughout this paper except where otherwise indicated.

Discounted for inflation, the average price paid for R01s has increased very slightly over the last decade, from $78,117 to $85,888; that for P01s has actually declined, from $570,377 to $470,268.[4] Examined over an even longer epoch — FY 1970 to FY 1988 — the average constant dollar size of an R01 award increased at an average annual rate of 1.3%, while that of a P01, after a sharp rise in the early 1970s, declined at an average annual rate of 1.3% between FY 1975 and FY 1988. If account were taken of the "negotiated reductions" effected in FY 1988, the constant dollar prices paid for average R01s and P01s would have been $86,574 and $476,803, respectively. These differences would affect the average growth rates cited only trivially. Thus, the pattern of the last decade is a perduring one.[4]

In view of the modest growth in size of these awards after discounting for inflation, the issue would not be worthy of extended analysis under ordinary circumstances. But because of the intense interest it holds for Congress and for the Inspector General, closer examination, focusing on the R01 awards, is in order. The R01 is the best known and understood as well as the principal instrument of support in the NIH portfolio. On the average, it provides partial salary support for one principal investigator (PI), one other doctoral-level investigator, and about two nondoctoral research personnel, a pattern that has persisted for a couple of decades.

Basis for Increases

Several factors probably contribute to the observed small real increase in the average size of awards for R01s.

Indirect Costs. A disproportionately rapid rise in indirect costs is one significant factor. Between FYs 1979 and 1988, the real size of the average R01 attributable to direct costs increased from $55,497 to $58,019, or 4.5%, while that due to indirect costs rose from $22,574 to $27,584, or 22.2%. Actually, in four of the last nine years, direct costs were below the FY 1979 base, while indirect costs rose steadily and substantially.[9]

Personnel Charges to Grants. Since about 65% of all funds budgeted in the average R01 award are for personnel, the NIH conducted a detailed study of salary changes in approved budgets over the period FY 1981-1985.[10] The number of PIs per award is, by definition, one; it did not change. Among PIs, the degree of participation rose for investigators with clinical doctorates, declined for those with academic doctorates, and declined in the aggregate, as reflected in the change in full-time equivalent personnel (FTEs); there was a small increase in the percentage of staff approved to receive compensation from the award and a substantial percentile increase in the amount of compensation approved. The pattern for non-PI doctoral personnel was different: the number of personnel declined but the average degree of participation increased, resulting in a modest increase in the number of FTEs; the percentage of staff approved to receive compensation increased substantially (that is, fewer worked for no pay), and the amount of compensation approved for payment increased modestly. The pattern for nondoctoral personnel did not follow either of these patterns: both the number of personnel and degree of participation decreased, as therefore did the number of FTEs; the number approved for compensation from the award increased slightly, as also did the approved level of compensation.

---

[*] For technical accounting reasons, data reported by the NIH on awards and on obligations differ slightly, as explained in the Definition section of the *NIH Data Book 1988*.[3] The time series describing obligations is one year more current than that for awards and so was used wherever data for FY 1988 and estimates for later years were cited.

Although the patterns of the number, commitment, and compensation of personnel varied by type of worker, the important net effect was that the real annual growth rate in compensation from the award increased only for the PIs: by 3.5% for those with clinical doctorates and by 0.9% for those with academic doctorates. For all other workers, the annual growth rate in compensation from the award declined: 2.3% per annum for non-PI personnel with doctorates and 3.3% per annum for nondoctoral personnel. In absolute terms, compensation from the grants, unadjusted for inflation, increased between FYs 1981 and 1985 from $12,420 to $16,648 for investigators with academic doctorates, while for those with clinical doctorates, compensation went from $17,797 to $26,353 over the same period.

Another NIH study of the aggregate personnel budgets of all R01 awards in 1980 and in 1985 showed a direct cost increase in the constant-dollar amounts approved, equivalent to annual growth rates of 0.5% for competing awards and 1.5% for noncompeting awards.[11]

These data suggest that salary increases in excess of inflation make only a small contribution to the observed constant-dollar increase in the size of the average R01.

The NIH analyses of salary changes are not entirely concordant with the conclusions drawn by the DHHS Inspector General in his March 1989 Preliminary Report to Congress,[12] based on an examination in eight institutions of a sample of grants. The differences between the NIH's and the Inspector General's studies include the time spans covered (1981–1985 versus 1984–1988), the item examined (budgeted versus paid salary), the types of awards analyzed (NIH R01s versus a mix of grants and contracts supported by various agencies), the inflator used (BRDPI versus CPI), and the field of science pursued (all biomedicine versus 50% non-biomedicine). The Inspector General's study needs refinement, extension, specification, and sophistication to be acceptable; the NIH's needs updating.

Quality of Research. A third factor possibly contributing to the small constant-dollar increase in the average award size relates to the quality of the research. The experience of the NIH is that the "better" the priority score accorded an application by study sections, the larger the size of the award approved and received.[4] The frequency distribution of priority scores has shifted during the last decade,[9] with a growing proportion of applications being deemed more meritorious, that is, with priority scores better than 150 or 175. Between FYs 1979 and 1988, the percentage of eligible applications with scores of between 100 and 150 rose from 10% to 27% of this total, while those with scores of between 100 and 175 increased from 22% to 41%. Even more striking is the representation in the award pool of applications accorded the highest technical merit. Over the same interval, awards with scores of between 100 and 150 soared from 20% to 71% of the total while those with scores of between 100 and 175 rose from 42% to 91%. The enhanced concentration in the award pool of the best—and most expensive—applications would tend to increase the average award size.

This exploration of possible bases for the observed growth in the average size of an R01 award should not distract from the central reality. Once inflation has been discounted, only a very small increase has occurred in the size of the traditional (R01) research award.

The Total Cost of the NIH Program

Another context in which the concerns about increasing costs are being articulated relates to the total annual federal investment in the NIH program. As noted earlier, this has increased steadily in current dollars and also, after a small fall between FYs 1979 and 1982, in constant dollars.

In the last decade, the number of R01 awards increased from 14,320 to 16,499, and the current-dollar expenditures, from $1.107 billion to $2.548 billion; the number of P01 awards grew from 456 to 746, while funding expanded from $253 million to $625 million. Thus, these two types of awards, which jointly accounted for about 85% of all research project expenditures in FY 1988, have expanded both in numbers and in obligations, whether the latter be in current or constant dollars.[4]

As previously noted, several new research project grant mechanisms have been created recently. Since new instruments of support generally expand at a rapid rate before eventually stabilizing, it is premature to discuss trends for them. But the absolute level of these awards is worth noting. The New Investigator and FIRST award programs are in the process of merging; together, they aggregated to almost 1,500 awards (1,472) in FY 1988 and accounted for $120 million in obligations. There were 598 MERIT awards and 71 Outstanding Investigator Awards during the same year. Compared with 16,500 R01s, the number of awards under the new mechanisms is small but not insignificant. Another characteristic of them is worth noting: they tend to carry a longer than average number of years of committed support.

Direct NIH investments in training have risen slightly in current dollars but fallen dramatically when discounted for inflation, continuing a trend that has been evident for almost two decades.[2,13] In FY 1969, about 2.5 times as many real dollars per year were invested in research training as was the case in FY 1988. Moreover, the number of FTE trainees has fallen significantly from the early-1970s level and the average real-dollar support per trainee is now at about 60% of the FY 1969 level.[2,13] Curiously, collateral evidence indicates that the first-year and total graduate enrollments of American students in the biosciences[14] as well as the number of Ph.D.s awarded to U.S. citizens[15] and the number of American postdoctoral students in these fields have been quite stable.[14] If, as many believe, the support of the training of graduate students has been, and continues to be, shifting imperceptively from training grants to research project grants, it is puzzling that the trend has not increased the constant dollar expenditures for an average R01 more extensively. Perhaps graduate students are paid less than technicians when the students replace them.

Clearly, the last decade has seen real and substantial growth in the overall extramural program of the NIH accompanied by some changes in the distribution of funds among various activities.[2] Indirect costs, though not the largest expenditure, have grown at the most rapid rate. The constant-dollar obligations for indirect costs, research projects, and intramural research were higher in FY 1988 than in FY 1978, those for research centers were about the same, while those for training and research contracts declined. The most dramatic relative expansion has been in the research project grant category, with other instruments of support—except for training and contract research—holding their own.

Expenditures for research related to the acquired immunodeficiency syndrome (AIDS) accounted for perhaps 20% of the increases over the decade.[16]

The Rising Current Services Requirements

Perhaps the most interesting problem is why the current services requirements for the NIH budget seem to have increased recently at a rate significantly more rapid than that of inflation. In the development of its annual budget request, every federal agency must submit an estimate of how large an appropriation it would need to maintain in the next fiscal year the level of program activity that is under way in the current fiscal year. This is called the current services requirement; as a first approximation, it should amount to the current year's appropriation plus an adjustment for inflation.

Estimates of current services requirements warrant more than casual discussion. Ever since enactment of Gramm-Rudman-Hollings, they have assumed increasing importance as Congress has sought to adjust available resources to the statutory limits imposed on the magnitude of the deficit. The several forms taken by perennial proposals for "freezes" contain implicit assumptions about current services requirements. Out-of-line estimates draw quick and intense scrutiny. As noted earlier, the NIH's total current services estimate for FY 1989 was 14% higher than the FY 1988 level, sufficiently beyond expectations to prompt Congress to ask the Inspector General to look into the matter. The same phenomenon occurred in 1989, when the NIH's current services estimate for FY 1990 required a 12% increase, again well in excess of inflation.[17] Since growth, positive or negative, has been modest during the last few years for other types of support mechanisms, the basis for the observed phenomenon resided in the research project grant category, which required an 18% increase.

At least three factors involving project grants could account for the observation—real growth, "negotiated reductions," and project period extension. These are difficult to tease apart and it may therefore be helpful to review—in idealized form and in isolation—the implications that each carries for project grant budgets. A small digression to explain how the NIH financing system operates may facilitate understanding these phenomena.

Throughout the past decade, the NIH has annually tendered about 4,500 three-year R01 awards; the project period for these awards averaged three years. The 4,500 competing first-year awards are also prebudgeted for future payment in the budgets for the next two years for the second and third years of support. By long-standing policy, only the first year award is obligated from the current year's appropriation; out-year noncompeting commitments have, however, first claim an out-year funds. Thus, in *any* steady-state year, there would be about 4,500 competing awards and about 9,000 noncompeting continuation ones, a total of about 13,500 awards to be paid (Fig. 1). Under these circumstances the current services budget for a steady-state year should reflect the costs of 4,500 competing awards and 9,000 noncompeting awards at their prior year's costs, adjusted for projected inflationary increases (Fig. 2).

Program Growth

If the number of competing awards were to grow from the steady-state level of 4,500 by 500 per year for three years (Fig. 3), this expansion in the *competing* pool would increase seriatim the size of the *noncompeting* pool for the next four years until a new steady state would be attained, two years after the number of competing awards stabilized at the new level of 6,000. In any year after growth had begun but before the new steady state had been reached, the current services requirements would exceed expectations by a considerable margin. Consider the current services requirements for the fourth year after the expansion began (Fig. 4). The expectation would be simply an inflationary escalation of the costs of 6,000 competing and 10,500 noncompeting awards funded in the prior year. The actual costs, however, must cover the inflation-escalated out-year costs of the earlier increases in competing awards, now annually expanding and showing up in the noncompeting pool.

As already indicated, the NIH's research project grant programs have experienced modest growth for the

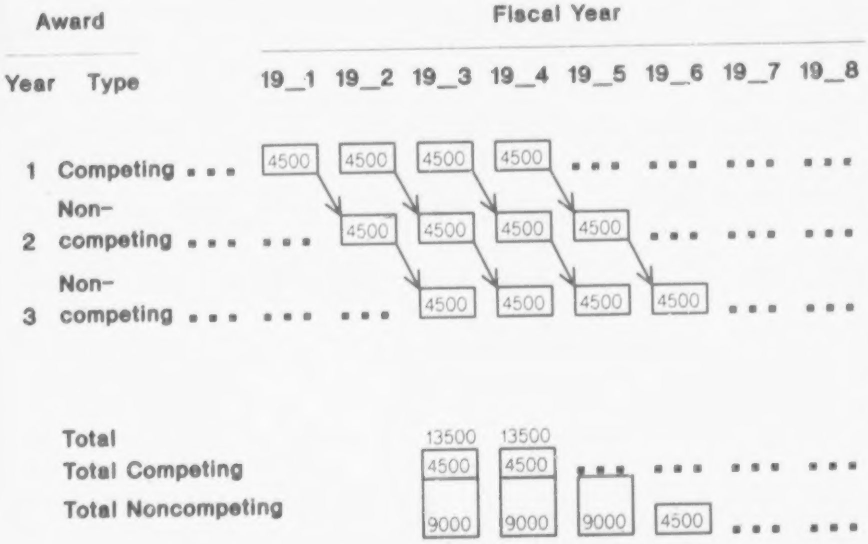


Figure 1. **An idealized snapshot in time of the NIH budgeting process, showing the effect, on the pool of noncompeting awards, of the competing awards made in FYs 19_1 through 19_4. The scheme depicts a steady-state situation in which 4,500 competing awards, each with a project period of three years, are made each year. The dots imply that the same process preceded or followed the events depicted. Each competing first-year award is funded from the appropriation for that year; the second and third (noncompeting) years of each award are funded in, and from the appropriations for, subsequent years. The number of awards in the noncompeting pool for any year is the sum of the number of competing awards made in the two previous years.**

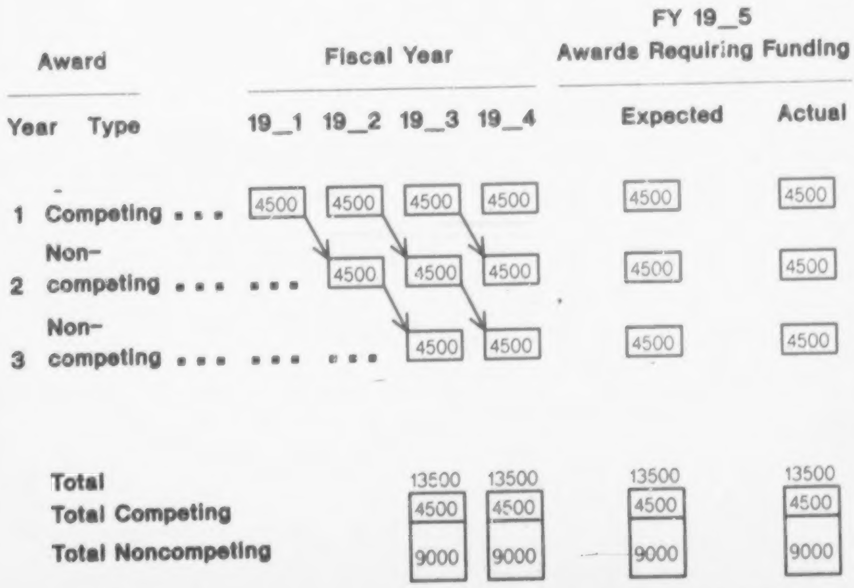


Figure 2. **The comparison, idealized, of actual and expected FY 19_5 current services requirements, i.e., the number of awards that must be funded (at inflation-adjusted costs) in a year for which a budget is submitted, e.g., FY 19_5. The number of competing awards is, by definition, the same number as that in the prior year; the number of noncompeting awards is the number to which commitments had been made in previous years. Under the assumptions specified in Figure 1, the expected and actual numbers coincide.**

last decade, a fact that explains, in part, why the current-services-requirements estimates have been unexpectedly high in the last few years.

"Negotiated Reductions"

The large "negotiated reductions" in award size that have been imposed in FYs 1986, 1988, and 1989 also distort current services estimates; such distortion is most easily and quickly apparent for competing awards. Regardless of the dynamics of the *noncompeting* pool, the number of competing awards in a current services projection is, by definition, exactly the same as that in the current year; therefore, the appropriations needed to fund that number should be identical in the two years except for an inflationary adjustment. However, in the spring of 1989, the NIH's FY 1990 current services projection for the same 5,324 competing awards was almost 17% higher than their FY 1989 costs.[17] Why? Recall that the FY 1989 competing awards had been arbitrarily reduced by 10.0%. Had the current services requirements been based on the amount approved by the study section, prior to the "negotiated reduction," the current services projection would have been just about equal to the FY 1989 costs plus inflation. The apparently disproportionally high current services estimate for competing awards is, in a sense, an artifact of "negotiated reductions."

"Negotiated reductions" are, of course, applied to the current-year amounts of both competing and noncompeting awards, although the magnitude of the cuts has been less for the latter than the former.[17] The current-services-requirements projections are distorted for both pools. However, other factors, such as real program growth and project period extension, may also be exerting influence on the noncompeting pool; thus, the effect is not as unambiguously attributable to one specific cause as is the case for competing awards.

Project Period Length

The third factor that can affect the current-services-requirements estimates is the length of the project period. For a long time, the average project period of R01 awards was very close to three years. Consider the effect of extending it from three to four years. Nothing would happen immediately after adopting a policy to extend project periods; the pools of competing, noncompeting, and total awards remain stable (Fig. 5). But in the third year after the first extended project period awards were tendered, there would be an abrupt increase in the noncompeting pool and a simultaneous rise in the current-services-requirements estimates, in excess of inflation, to cover the expanded number of noncompeting continuations (Fig. 6).

What has actually been happening to the average project periods of NIH awards? For the new types of research project awards that originated in the 1980s, such as the New Investigator and FIRST awards, five-year project periods were the norm; Outstanding Investigator awards were of even longer duration. Starting in FY 1985, the number of research project grants with three-year project periods began to decline and that of those with five-year periods to rise, for the eligible grants (that is, those that were approved) but also, especially, for those actually awarded.[9] By FY 1988, the incidence of five-year awards exceeded 50%, up from less than 20% in FY 1984.

Looked at in somewhat more detail (Fig. 7), a significant number of five-year-program project (P01) awards have been made throughout the decade, but in recent years the fraction of these has increased abruptly while that of three-year awards has simultaneously fallen. For R01s, a significant redistribution of project period lengths of awards began in FY 1985 and continues.[4] Stated another way, the average lifespan of an R01 award increased slightly between 1979 and 1980 but increased substantially after FY 1985, from 3.30 years to 3.92 years in FY 1988.[2]

To project the impact of extending project periods, the expected size of the noncompeting R01 pool can easily be calculated using two items of actual data: the number of competing awards for a given year and the frequency distributions of their project periods. To test for validity, the calculated and observed numbers in the noncompeting pool from FY 1983 through FY 1988 were compared; the calculated numbers varied less than 1% from the actual ones, indicating

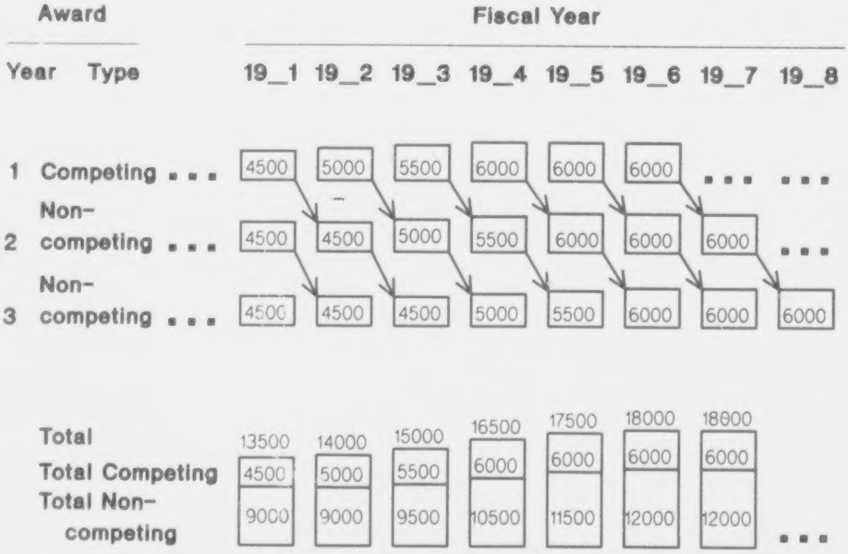

Award
Fiscal Year
Year
Type
19_1
19_2
19_3
19_4
19_5
19_6
19_7
19_8
1
Competing . . .
4500
5000
5500
6000
6000
6000
. . .
. . .
2
Non-competing . . .
4500
4500
5000
5500
6000
6000
6000
. . .
3
Non-competing . . .
4500
4500
4500
5000
5500
6000
6000
6000
Total
Total Competing
Total Non-competing
13500
4500
9000
14000
5000
9000
15000
5500
9500
16500
6000
10500
17500
6000
11500
18000
6000
12000
18000
6000
12000
. . .

Figure 3. **An idealized snapshot in time of the effect on the noncompeting award pool of (1) increasing the number of competing awards from the steady-state level that prevailed before FY 19_1 by 500 per year through FY 19_4, and (2) thereafter stabilizing that number at the new (6,000 per year) level. All awards are assumed to have three-year project periods. The size of the noncompeting pool does not stabilize until two years after growth in the competing pool ends.**

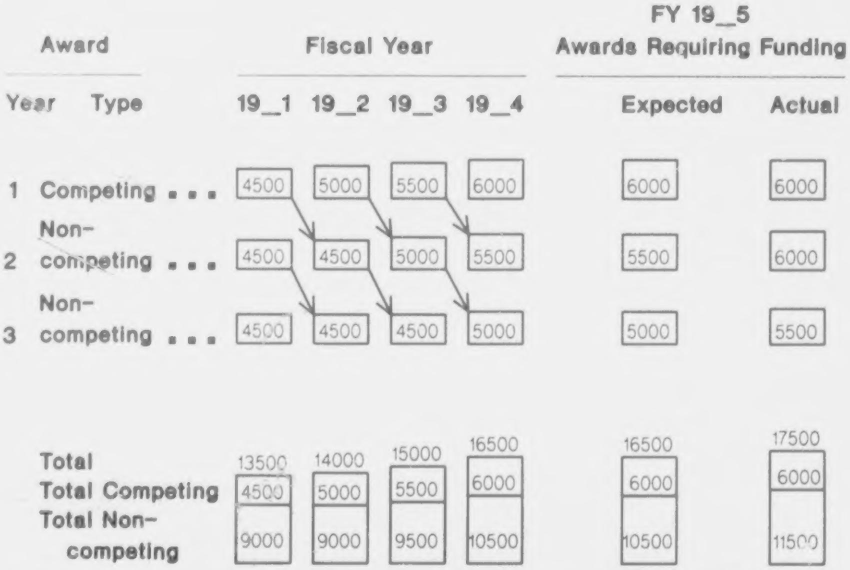

Award
Fiscal Year
FY 19_5
Awards Requiring Funding
Year
Type
19_1
19_2
19_3
19_4
Expected
Actual
1
Competing . . .
4500
5000
5500
6000
6000
6000
2
Non-competing . . .
4500
4500
5000
5500
5500
6000
3
Non-competing . . .
4500
4500
4500
5000
5000
5500
Total
Total Competing
Total Non-competing
13500
4500
9000
14000
5000
9000
15000
5500
9500
16500
6000
10500
16500
6000
10500
17500
6000
11500

Figure 4. **The comparison, idealized, of the actual and expected FY 19_5 current services requirements, i.e., the number of awards that must be funded (at inflation-adjusted costs) in a year for which a budget is submitted, e.g., FY 19_5, under the circumstances of program growth assumed in Figure 3. The actual number exceeds the expected because the former must include the increased number of commitments for third-year support that were occasioned by the increased number of FY 19_3 competing awards, as well as the increased number of commitments for second-year support occasioned by the increased number of FY 19_4 competing awards.**

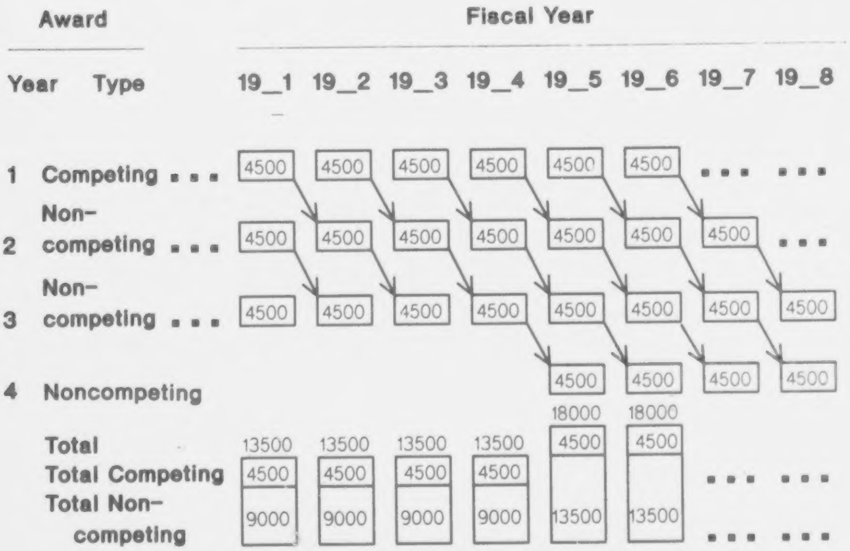

Award
Fiscal Year
Year Type
19_1 19_2 19_3 19_4 19_5 19_6 19_7 19_8
1 Competing . . .
4500 4500 4500 4500 4500 4500 . . . . . .
2 Non-competing . . .
4500 4500 4500 4500 4500 4500 4500 . . .
3 Non-competing . . .
4500 4500 4500 4500 4500 4500 4500 4500
4 Noncompeting
4500 4500 4500 4500
Total
13500 13500 13500 13500 18000 18000
Total Competing
4500 4500 4500 4500 4500 4500 . . . . . .
Total Non-competing
9000 9000 9000 9000 13500 13500 . . . . . .

Figure 5. **An idealized snapshot in time illustrating the effect upon the noncompeting award pool of making only one change: extending the duration of project periods from three to four years, beginning in FY 19_2. No impact becomes visible until FY 19_5, when the first cohort of awards with four-year project periods appear in the noncompeting pool. Thereafter, the new steady state is characterized by a 50% increase in the size of the noncompeting pool and a 33% increase in the total pool.**

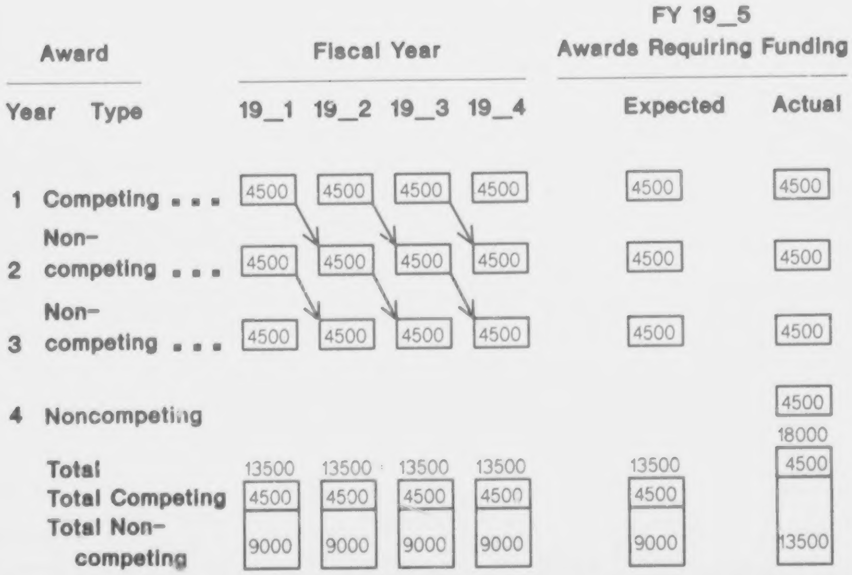

Award
Fiscal Year
FY 19_5
Awards Requiring Funding
Year Type
19_1 19_2 19_3 19_4
Expected
Actual
1 Competing . . .
4500 4500 4500 4500 4500 4500
2 Non-competing . . .
4500 4500 4500 4500 4500 4500
3 Non-competing . . .
4500 4500 4500 4500 4500 4500
4 Noncompeting
4500
Total
13500 13500 13500 13500 13500 18000
Total Competing
4500 4500 4500 4500 4500 4500
Total Non-competing
9000 9000 9000 9000 9000 13500

Figure 6. **The comparison, idealized, of the actual and expected FY 19_5 current services requirements, i.e., the number of awards that must be funded in a year for which a budget is submitted, e.g., FY 19_5, under the circumstances of project period extensions assumed in Figure 5. The actual number exceeds the expected because it must include the first cohort of extended-project-period awards that were made in FY 19_2.**

the reliability of the technique (Table 2). Obviously, actual award numbers for FY 1989 and beyond are not available because data for FY 1989, just completed, are not yet compiled. However, assuming for FY 1989 and thereafter a level of competing awards equal to that in FY 1988 and assuming no further change in the frequency distribution of project period durations from that which obtained in FY 1988, the computed size of the noncompeting pool rises through 1992. If the proportion of longer awards were to continue to increase, the noncompeting pool would increase even more rapidly and continue to do so beyond FY 1992.

Another interesting calculation compares the actual (up to FY 1988) and projected (after FY 1988) sizes of the noncompeting pool that resulted from the lengthening of project periods with what they would have been if the distribution of project periods had been frozen in the FY 1984 pattern. Results (Table 2) indicate that the practice of extending the duration of project periods begun in FY 1985 has had an effect, delayed as expected until FY 1988, but promising eventually to reach major dimensions even if the number of competing R01s remains stable or decreases. As can also be inferred, the practice of extending project periods has placed a substantial demand on the FY 1990 appropriation now under discussion in Congress and will place even greater demands, no matter what else happens, on appropriations for FYs 1991 and 1992. Depending on how many competing awards can be made, it could become necessary to find funds for as many as 2,750 "extra" R01s in FY 1992.

The forces at play on R01 awards are also operating on P01s. Moreover, most of the new programs recently included in the research project grant category of the budget are characterized by long project periods.

Discussion

The relative stability in the average size of R01 awards is not widely recognized. It is remarkable, however, that over an era in which science has become much more complex, instruments and equipment much more costly, and competition from industry

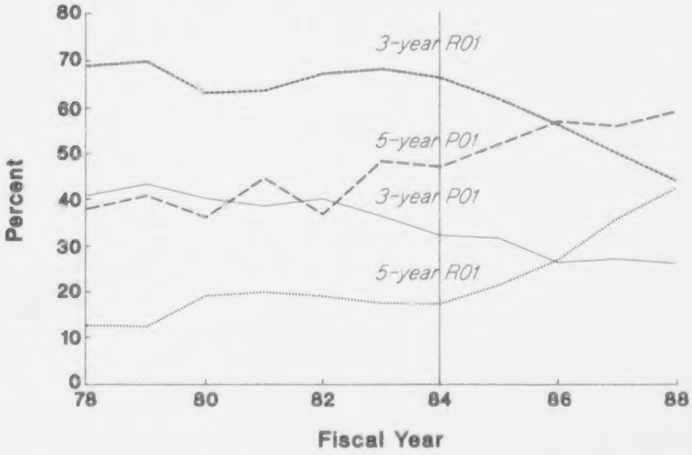


Figure 7. **Distribution of project period lengths of R01 and P01 awards. The percentages of all R01 and P01 awards assigned either three-year or five-year project periods have been plotted. During FY 1984, NIH officials openly encouraged advisory bodies to provide longer support to projects with high scientific merit. Thus, the percentage of both R01 and P01 awards with five-year project periods rose, with a corresponding decline in the percentage of awards with three-year project periods.**

and medical practice for academic talent increasingly intense, R01 award size has increased only at an annual rate of about 1.1–1.3% (depending on the epoch examined) in real dollars. This suggests that significant gains in efficiency and productivity must have been achieved or that funding from other sources has supplemented the federal award in the performance of the scope of work for which the latter was tendered.

Real growth in the size of the NIH program over the decade is somewhat unexpected, in view of the fiscal restraints that have been operative. However, the answer to the question of why NIH-funded biomedical research program costs the government more is simple. There is more of it!

It is difficult to isolate the effect of any one of the phenomena that increase current services requirements, but all are probably contributing, or will soon contribute, to the dis-ease surrounding the NIH budget.

Growth in the number of competing awards, whether steady or episodic, expands the commitment base and results in current services requirements that regularly exceeds what would be expected from inflation alone. The so-called "negotiated reductions" do not alter the fundamental financial resources required to complete a project's scope of work. To the extent that the money "saved" is used to sustain or increase the num-

Table 2

Calculations of the Effect of the Extensions of R01 Project Period Durations on the Numbers of Noncompeting R01 Awards for Fiscal Years 1983–1988 Compared with the Numbers Expected if Project Period Durations Had Remained at the 1984 Level, with Projections for Fiscal Years 1989–1992 That Involve Explicit Assumptions about the Number of Competing Awards and Project Period Durations

	Competing	Noncompeting		
Year	Observed Number	Observed Number	Calculated Number*	Expected Number†
1983	4,737	10,428	10,252	NA
1984	4,122	10,692	10,728	NA
1985	5.209	10,896	10,913	10,890
1986	4,966	11,243	11,113	11,282
1987	5,026	11,501	11,439	11,516
1988	4,651	11,842	11,713	11,543

	Projections			
Year	Assumed Number	Observed Number	Calculated Number‡	Expected Number§
1989	4,651	NA	12,171	11,267
1990	4,651	NA	12,720	10,894
1991	4,651	NA	13,357	10,815
1992	4,651	NA	13,524	10,774

*The values tabulated were calculated from the actual number of competing awards made and the actual distribution of project period durations assigned for each year from FY 1978 to FY 1988.

†The values tabulated for FY 1985 to FY 1988 were calculated from (1) the actual number of competing awards for each year from FY 1978 to FY 1988; (2) the actual distribution of project period durations for each year from FY 1978 to FY 1984; and (3) the FY 1984 distribution of project period durations.

‡The values tabulated were calculated using a number of competing awards for each year from FY 1989 to FY 1992 and a distribution of project period durations for those same years, each of which was equal to the values of those variables in FY 1988.

§The values tabulated were calculated using a number of awards for each year from FY 1989 to FY 1992 equal to that for FY 1988, and a distribution of project period durations for each of those years equal to the distribution that obtained in FY 1984.

ber of competing awards, it inevitably results in an expansion in the commitment base and in unexpectedly high current services requirements for both the competing and the noncompeting grants in the following year. The consequences of project-period extension vary, depending upon concomitant changes in the number of competing awards. Extension per se (that is, going from an average project period of three years — with 4,500 competing awards, 9,000 noncompeting ones, and a total of 13,500 — to an average project period of five years with the same total number of awards) would result in a new steady state, characterized by 2,700 competing and 10,800 noncompeting awards. The increased current services requirements attributable to the expanded noncompeting pool would be offset by the reduced number of competing awards. However, if the transition were to be accompanied by stability in the number of competing awards, the new steady state would be characterized by 18,000 noncompeting awards and a total of 22,500 awards. In this case, the current services requirements, unadjusted to inflation, would rise by 67%. The impact of real growth, occasioned by sustaining the number of competing awards at or near the pre-extension level, would, whatever its magnitude, be slow to appear and slow to respond to control measures.

Project period extension also *reduces award rates* by delaying the interval before a noncompeting award and the funds associated with it reenter the competing pool. Since, on the average, noncompeting awards remaining in that pool would have had award rates of about 55%, had they reentered competition, the nominal award rate is heavily weighted toward that for new competing awards, that is, about 25%. Thus, during the transition years in which project period length is being systematically changed, nominal award rates reflect a reality that is substantially different from that in the steady state.

The greatest concern would seem to be the impact of project period extension, because unless the number of competing awards declines, it will inexorably increase the size of the noncompeting pool over the next several years, and therefore increase funding requirements, precisely coinciding with reduced growth in real federal resources and ever keener competition for them. While lengthening project periods, at least selectively, is a highly desirable goal, adoption of a policy to do it is not without its price. Unless there is a reduction in the number of competing awards, the magnitude of downward "negotiations" will be virtually inescapable, unless there are large increases in appropriations.

Conclusions

During the last decade the total appropriations for the NIH have grown in current as well as constant dollars. Constant-dollar expenditures for indirect costs and research project grants have increased, as also has the number of the latter, while constant-dollar expenditures for research centers, training, and research contracts have shrunk. The most impressive redistribution in emphasis has been toward research project grants.

The size of the average R01 award, discounted for inflation, has grown at an annual rate of 1.1% during the last decade and 1.3% since FY 1970; that of the average P01 had declined over the same periods, after a slight rise in the early 1970s.

The modest rise in the real (constant-dollar) size of the average R01 reflects, inter alia, disproportionately rapid increases in indirect as compared with direct costs, growth in excess of inflation in compensation to PIs, especially those holding clinical doctoral degrees, and increased density in the pool of awards of applications accorded the most meritorious scores by the peer review system.

The regularity with which current-services-requirements estimates for the NIH exceed inflation reflects real growth in the program, particularly in the category of research project grants; the artifact of basing calculations on the post- rather than pre-"negotiated" levels of awards in the "current" year; and the extent to which the project periods of awards have been extended. The effect of lengthening project periods is slow to appear, but unless accompanied by the award of fewer competing applications, inexorably swells the pool of noncompeting awards; decisions in this area undertaken in 1985 and continued at least through FY 1988 will very significantly increase current services requirements in FYs 1991 and 1992.

Members of the NIH staff have been immensely helpful and responsive to requests for data. Particular thanks are due Robert F. Moore and Nadine Sutton, Information Systems Branch, DRG, Dr. William E. Rhode, Dr. Patricia E. McKinley, and Dr. Charles R. Sherman of the Office of Science Policy and Legislation; and H. Richard Miller and Anne Sumner of the Division of Financial Management. Thanks also to my colleague, David M. Moore, for the calculations in Table 2.

References

1. U.S. Congress. Congressional Budget Office. *The Economic and Budget Outlook: Fiscal Years 1990–1994.* Appendix F. Washington, D.C.: Government Printing Office, 1989.
2. National Institutes of Health. Office of Science Policy and Legislation. Division of Research Grants. *NIH Data Book 1988. Basic Data Relating to the National Institutes of Health.* NIH Publication No. 89-1261. Washington, D.C.: NIH, December 1988.
3. U.S. Congress. House. Committee on Appropriations. *Department of Labor, Health and Human Services, and Education, and Related Agencies Appropriation Bill, 1989.* Report 100-689. June 10, 1988. H. Doc. 85-653.
4. Personal communication. National Institutes of Health, Division of Research Grants, Information Systems Branch, 1989.
5. U.S. Department of Commerce. Bureau of Economic Analysis. *Survey of Current Business.* Washington, D.C.: Government Printing Office, September 1989.
6. U.S. Department of Labor. Bureau of Labor Statistics. *CPI Detailed Report, August 1989.* Washington, D.C.: Government Printing Office, October 1989.
7. U.S. Department of Commerce. Bureau of Economic Analysis, Government Division. *The Biomedical Research and Development Price Index (BRDPI).* Revised and updated BRDPI. Washington, D.C.: [unpublished report] December 1986.
8. National Institutes of Health. Office of Science Policy and Legislation. *Correction and Update of the Biomedical Research and Development Price Index (BRDPI).* Washington, D.C.: [memorandum] December 12, 1988.
9. National Institutes of Health. *Extramural Trends FY 1979–1988* Washington, D.C.: Government Printing Office, June 1989.
10. National Institutes of Health. Office of

Program Planning and Evaluation. *Changes in Research Project Participation and Compensation Patterns in Competing R01 Awards to Institutions of Higher Education (1981-1985).* Program Analysis Note No. 11. [Memorandum] July 21, 1986.

11. Personal communication. National Institutes of Health, Office of Program Planning and Evaluation, 1989.
12. U.S. Department of Health and Human Services. Office of the Inspector General. Office of Audit. *Survey Report on the Cost of Research at Colleges and Universities.* Common I.D. Number A-12-89-00128. Washington, D.C.: USDHHS, March 1989.
13. National Institutes of Health. Office of Program Planning and Evaluation. Division of Research Grants. *Basic Data Relating to the National Institutes of Health, 1976.* Washington, D.C.: NIH, February 1976.
14. National Science Foundation. Science and Engineering Sector Studies Group. *Selected Data on Graduate Science/Engineering Students and Postdoctorates, Fall 1988.* Washington, D.C.: National Science Foundation, September 1989.
15. National Research Council. *Summary Report 1987: Doctorate Recipients from United States Universities.* Washington, D.C.: National Academy Press, 1989.
16. U.S. Congress. House. Subcommittee on the Departments of Labor, Health and Human Services, Education, and Related Agencies. *Hearings before a Subcommittee of the Committee on Appropriations, House of Representatives, One Hundredth Congress, Second Session.* Part 4-A, National Institutes of Health. 1988. H. Doc. 84-506.
17. Personal communication. National Institutes of Health, Office of the Associate Director for Administration, Financial Management Branch, 1989.

ROBERT G. PETERSDORF, M.D.

Three Easy Pieces

Abstract—The author provides information about and analyzes three issues confronting academic medical centers in the realms of education, patient care, and research. (1) In the educational realm, he indicates why medical centers must play an expanded role in training primary care physicians, explains the dangers of not doing so, and describes ongoing and proposed approaches and reforms for achieving this goal. (2) In the arena of patient care, he explains why modifying physician reimbursement policies is essential for more physicians to develop careers as generalists. Other more controversial physician payment reform measures and their implications for health care and academic medical centers are discussed; the author urges that benefit to patients always be the first concern of such reforms, even at the expense of more narrowly based interests such as limits in faculty salaries or reduced overages to institutions. (3) Regarding research, he discusses various facets and implications of conflict of interest for biomedical scientists—both the reality of misconduct and the appearance of it—especially as they apply to the growing number and forms of university-industry relationships, and urges that such conflicts be managed within guidelines that clarify expectations and standards in an atmosphere of appropriate disclosure and oversight. He concludes by urging academic medical centers to rise to, rather than avoid, the three challenges he has described. *Acad. Med.* 65(1990):73–77.

In 1970, Jack Nicholson starred in the critically acclaimed film "Five Easy Pieces." Nicholson, in an important lead role that helped to establish his career, played a brilliant classical pianist from an eccentric musical family. He had abandoned his family's rich musical tradition in favor of the California oil fields, and the film was seen as an exposition of the rootlessness of contemporary American life.

The film's title, deemed provocative and mysterious by many reviewers, actually refers to a practice book used by most youngsters studying the piano.[1] I thought it might be a good metaphor for my remarks, which are meant to be a guide to some issues confronting the academic medical center. In the interest of time, I have discounted the five easy pieces to three, the better to coincide with the education, patient care, and research missions of academic medical centers.

Education

The first of my easy pieces is entitled "Primary Care Overture." Many of you have heard me play this piece before; it is a standard opus in my repertoire. Like Rossini's famous "William Tell Overture," my overture is in danger of becoming known more familiarly as the theme from the "Lone Ranger." Yet I persist in my efforts to seek converts among the academic medical community because I am convinced that our agenda must encompass an expanded role for primary care in our institutions and must include an expanded role in training primary care doctors for our people.

In the 1960s the federal government challenged the academic medical community to expand its capacity to address a perceived shortage of physicians. The response was vigor-

---

This paper was presented at the opening plenary session of the 100th Annual Meeting of the Association of American Medical Colleges, held in Washington, D.C., October 1989.

Dr. Petersdorf is president, Association of American Medical Colleges, Washington, D.C. Correspondence may be addressed to him at AAMC, One Dupont Circle, N.W., Suite 200, Washington, D.C. 20036. Reprints will not be available.

RESEARCH REPORTS

ARTICLES

Comparing Research Activities of Women and Men Faculty in Departments of Internal Medicine

BARBARA A. LEVEY, M.D., NANCY O. GENTILE, H. PAUL JOLLY, Ph.D.,
HARRY N. BEATY, M.D., and GERALD S. LEVEY, M.D.

Abstract—The findings from a questionnaire prepared by the Association of Professors of Medicine and the Association of American Medical Colleges were published in two reports in 1986 and 1987. These reports assessed the research activities of full-time members of departments of internal medicine in 1982 and 1983. The purpose of the present study was to analyze the data of the earlier reports in order to compare the research activities of women and men who were full-time faculty in departments of medicine during the time period originally surveyed. More than half of the faculty women who responded (52%) were less than 40 years old, compared with 23% of the faculty men. Sixty-seven percent of the women held the rank of instructor or assistant professor, in contrast to 40% of the men holding these ranks. Although the faculty of both genders reported generally the same proportions of time devoted to research, the women researchers with M.D. degrees had significantly less National Institutes of Health (NIH) grant support than did their counterparts who were men. Since this difference may have been a function of age, the authors compared NIH grant support of the faculty men and women with M.D. degrees who were 40–59 years old. Even in this older group, significantly fewer of the faculty women had NIH grant support than did the men (16% versus 30%). Furthermore, the percentage of designated laboratory space was significantly lower among the faculty women, regardless of degree (M.D., M.D./Ph.D., or Ph.D.). Further investigation is warranted to monitor the progress of women attempting to develop their research careers and to assess their overall clinical teaching and administrative roles in departments of medicine. *Acad. Med.* 65(1990):102–106.

Discussions of the progress of women in medicine, and more specifically, in academic medicine, appear cyclically in the scientific literature.[1–5] Relman[3] stated in 1980 in an editorial, "Here Come the Women," that internal medicine was the favorite specialty among women medical school graduates. The following year Braslow and Heins,[4] in a comprehensive paper, "Women in Medical Education," expressed concern for the underrepresentation of women at the level of full professors. Koshland,[2] in a 1988 editorial published in *Science*, commented on the relative underrepresentation of women in the higher echelons of academia and industry. More recently, Bickel[5] published a status report on women in medical education in *The New England Journal of Medicine*, writing that the profile of women in medical education looks much as it did at the beginning of the decade and noting that faculty promotion policies strongly favor productive researchers. Bickel referred to unpublished data regarding research efforts of women in departments of medicine and also extracted data from two reports[6,7] that focused on the extent of research activities of full-time faculties in academic departments of internal medicine. Neither of these reports addressed specifically the degree of involvement of women who were internal medicine faculty in research. The purpose of the present study was to analyze the data from the survey on which these two earlier reports were based in order to compare the research activities the women and men who were full-time faculty in academic departments of medicine during the period covered by the survey, 1982 and 1983. These were years in which the number of women physicians was increasing at a greater rate than that of men physicians.[8]

Method

The method of the 1982–1983 survey has been described[6]; a brief summary follows.

The survey used data from the Fac-

Dr. Barbara Levey is associate dean and director of admissions, and also associate professor of medicine and pharmacology, Department of Medicine; Dr.Gerald Levey is professor and chairman, Department of Medicine; both at the University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania. Ms. Gentile was formerly Director, Faculty Roster System; Dr. Jolly is associate vice president for Operational Studies; both at the Association of American Medical Colleges, Washington, D.C. Dr. Beaty is dean, Northwestern University Medical School, Chicago, Illinois.

Correspondence and requests for reprints should be addressed to Dr. Barbara A. Levey, Associate Dean and Director of Admissions, University of Pittsburgh School of Medicine, M-245 Scaife Hall, 3550 Terrace Street, Pittsburgh, PA 15261.

ulty Roster of the Association of American Medical Colleges (AAMC). The Faculty Roster is a computer-based system containing data on full-time faculty of accredited medical schools in the United States. At the time a faculty member is first appointed, a Faculty Roster form is completed either by the new faculty member or by an administrator of the medical school. The file is updated by the medical school after each promotion, granting of tenure, transfer to another institution, or termination of the faculty member. The form includes questions in five groups, providing background information, data about current appointments, professional employment history, and information about education and training. It does not contain detailed information about research activities.

Table 1

Numbers and Percentages, by Age Group and Academic Rank, of 7,015 Men and Women Faculty in Internal Medicine Departments, 1982–1983*

	Men		Women	
Category	No.	%	No.	%
Age group†				
Younger than 40	1,615	23	485	52
40–59	4,533	65	387	42
60 or older	856	12	58	6
Missing	11	—	2	—
Total	7,015	100	932	100
Academic Rank‡				
Instructor	282	4	130	14
Assistant Professor	2,522	36	493	53
Associate Professor	1,881	27	183	20
Professor	2,252	32	99	11
Other and missing	78	1	27	3
Total	7,015	100	932	101

*These faculty were among the 7,947 who responded to a survey requesting them to verify their AAMC Faculty Roster record and to complete a research activity form from the Association of Professors of Medicine. (Complete data were not available for some respondents.) A total of 11,293 faculty were sent the survey materials.

†Chi-square for age group and gender = 358.74; $p < .0001$.

‡Chi-square for academic rank and gender = 332.2; $p < .0001$.

The 1982–1983 survey also used data from a separate research activity form devised by members of the Association of Professors of Medicine Task Force on Manpower Needs with the assistance of staff members of the Association of American Medical Colleges. Research activity was defined as "an activity performed with the objective of developing knowledge which usually leads to publication and which may be the basis for funding support." Questions were designed to ascertain research effort and sources of funding in the year before the 1982–1983 survey, as well as two, five, and ten years earlier; previous research training, use of research space, and publications that communicate research results.

A record outlining demographic and current appointment information was generated from the Faculty Roster system for each full-time, salaried faculty member previously listed in the system as a member of the department of internal medicine at an accredited medical school. All the Faculty Roster records and blank research activity forms were distributed to the chairmen of the departments of internal medicine in all U.S. medical schools. They distributed these materials to the faculty members in their departments and asked them to complete, verify, or correct the Faculty Roster records and to fill out the research activity forms. When this had been done, all the completed research forms and verified roster records were forwarded to the Faculty Roster at the AAMC.

Information included in the Faculty Roster database was verified or updated from the materials received, and the faculty data that had not been in the system were entered.

Survey Compliance

There were 11,293 Faculty Roster records generated for the survey; these and the same number of blank research activity forms were sent to departments of internal medicine at U.S. medical schools in the manner previously described. Complete data that could be used in the present study were returned by 7,947 faculty members. At the time of this survey there were 1,346 women in departments of internal medicine, representing 11.9% of the 11,293 faculty members who were sent the survey materials. A total of 932 women responded to the survey, which is 11.7% of the 7,947 faculty who responded. The numbers of responses to individual questions varied; that is, not all faculty members answered all the questions. For all the variables characterizing faculty members, there were no significant differences between the group of faculty who responded and the total group of full-time faculty to whom the Faculty Roster and research activity forms were sent.

Method of Analysis

As the research activity forms were returned, they were edited for accuracy, and missing answers were distinguished from zero answers. These data were merged with the full, verified record of the faculty member in the Faculty Roster database, using the member's Social Security number as the identifier. A more detailed description of these analytic procedures can be found in an earlier publication.[6]

Results

Analysis of the distribution of respondents by age and gender (Table 1) shows that the largest percentage of the women faculty (52%) were less than 40 years old, compared with 23% of the men in this group. In contrast,

Table 2

Numbers and Percentages, by Percentage of Effort Spent in Research, of 7,793 Men and Women Faculty with Either M.D., M.D./Ph.D., or Ph.D. Degrees Who Were in Internal Medicine Departments, 1982–1983*

% Effort in Research	Faculty with M.D. Degrees				Faculty with M.D./Ph.D. Degrees				Faculty with Ph.D. Degrees			
	Men		Women		Men		Women		Men		Women	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
0	1,747	29	206	32	79	16	7	17	28	6	14	8
1–20	1,473	25	182	28	84	16	9	23	18	4	5	3
21–40	956	16	90	14	110	22	10	25	21	4	5	3
41–60	790	13	68	10	92	18	7	17	30	6	8	5
61–80	585	10	60	9	85	17	6	15	119	25	16	10
81–100	226	4	30	5	49	10	1	3	261	54	112	68
Missing	174	3	13	2	8	1	0	0	4	1	5	3
Total	5,951	100	649	100	507	100	40	100	481	100	165	100

*These faculty were among the 7,947 who responded to a survey requesting them to verify their AAMC Faculty Roster record and to complete a research activity form from the Association of Professors of Medicine. (Complete data were not available for some respondents.) A total of 11,293 faculty were sent the survey materials.

the highest percentage of men was in the 40–59-year age group. The difference between the ages of the women and the men was highly significant ($p < .0001$). Table 1 also shows the distribution of men and women faculty by academic rank. Sixty-seven percent of the women had the rank of instructor or assistant professor; 40% of the men held these ranks. Thirty-one percent of the women faculty and 59% of the men held the rank of professor or associate professor. These data are highly significant ($p < .0001$), indicating that the women were more likely than the men to be junior faculty.

A preponderance of both the women (68%) and the men (85%) held the M.D. degree only. The percentage of the men holding M.D./Ph.D. degrees (7%) exceeded that of the women with such degrees. It is notable that 18% of the women held the Ph.D. degree only; 7% of the men held this degree only. These differences were also reflected in the total populations of men and women faculty in departments of internal medicine.

The results shown in Table 2 indicate a general agreement of the percentages of the men and the women engaged in research activities, regardless of degrees earned. For example, approximately 74% of the women who had M.D. degrees and 70% of the men with M.D. degrees reported spending 40% or less of their time in research. Almost one-third of all the men and women with M.D. degrees reported spending no time in research activity. Although the numbers of women and men who had Ph.D. degrees was small, it is evident that these faculty members spent a great proportion of their time in research; 78% of the women and 79% of the men reported greater than 60% of their effort was in research; only 18% of the women and 27% of the men with M.D./Ph.D. degrees reported that level of research effort. Similarly, a greater percentage of respondents with M.D./Ph.D. degrees than those with Ph.D. degrees reported no research effort.

Table 3 shows that the percentage of men researchers with M.D. degrees who had National Institute of Health (NIH) funding was roughly equivalent to the percentage of men researchers with either M.D./Ph.D. or Ph.D. degrees who had NIH grant support. In contrast, a much smaller percentage of women researchers

Table 3

Numbers and Percentages of 7,793 Men and Women Faculty in Internal Medicine Departments, Sources of Funding, and Academic Degrees Earned, 1982–1983*

Funding	Faculty with M.D. Degrees				Faculty with M.D./Ph.D. Degrees				Faculty with Ph.D. Degrees			
	Men		Women		Men		Women		Men		Women	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
NIH†	2,227	37	83	13	183	36	13	32	193	40	65	39
Other‡	1,733	29	283	44	228	45	17	43	241	50	81	49
None	1,991	34	283	44	96	19	10	25	47	10	19	12
Total	5,951	100	649	101	507	100	40	100	481	100	165	100

*These faculty were among the 7,947 who responded to a survey requesting them to verify their AAMC Faculty Roster record and to complete a research activity form from the Association of Professors of Medicine. (Complete data were not available for some respondents.) A total of 11,293 faculty were sent the survey materials.

†This group includes faculty who had other sources of funding in addition to their National Institutes of Health funding.

‡The faculty members in this group had internal and external funding from sources other than the NIH, such as the Veterans Administration, the National Science Foundation, and industry.

Table 4

Numbers and Percentages, by Source of Funding and Academic Rank, of 3,779 Men and Women Faculty, Aged 40–59 Years, with M.D. Degrees Who Were in Internal Medicine Departments, 1982–1983*

	Faculty with M.D. Degrees			
	Men		Women	
Category	No.	%	No.	%
Funding†				
NIH‡	1,023	29	45	16
Other§	1,478	42	126	44
None	992	28	115	40
Total	3,493	99	286	100
Academic rank‖				
Instructor	28	1	12	4
Assistant professor	622	17	112	39
Associate professor	1,349	38	106	37
Professor	1,474	42	51	18
Other and missing	20	1	5	2
Total	3,493	100	286	100

*These faculty were among the 7,947 who responded to a survey requesting them to verify their AAMC Faculty Roster record and to complete a research activity form from the Association of Professors of Medicine. (Complete data were not available for some respondents.) A total of 11,293 faculty were sent the survey materials.

†Chi-square for source of funding and gender = 29.92; $p < .0001$.

‡This group includes faculty who had other sources of funding in addition to their National Institutes of Health funding.

§The faculty members in this group had internal and external funding from sources other than the NIH, such as the Veterans Administration, the National Science Foundation, and industry.

‖Chi-square for academic rank and gender = 130.76; $p < .0001$.

with M.D. degrees than those with either M.D./Ph.D. or Ph.D. degrees had NIH grant support. Furthermore, the women scientists with M.D. degrees were much more dependent on other external (non-NIH) grant support than were their counterparts among the faculty men. The percentages of both the men and women faculty with M.D. degrees who reported no research funds of any type was high (44% of the women and 34% of the men).

Because of the relative youthfulness of the women in the population studied and the striking difference between the men and the women with M.D. degrees with regard to NIH funding, the authors examined all the faculty members with M.D. degrees who were 40–59 years old with regard to NIH funding (Table 4). Even in this more mature group of faculty, the women reported significantly less NIH funding than did the men. Table 4 also provides an examination of rank in this 40–59-year-old age group. The data show that even in the older age group, the women were significantly more likely than the men to be junior faculty ($p < .0001$).

Additional data not presented in the tables show that a greater percentage of the men than the women faculty reported assigned laboratory space regardless of degree. The percentages of the men versus the women faculty with assigned laboratory space were the following: 50% versus 32% of those with M.D. degrees, 71% versus 65% of those with the M.D./Ph.D. degrees, and 80% versus 67% of those with Ph.D. degrees. In addition, more women than men reported no assigned laboratory space.

Productivity of faculty researchers was assessed by the number of original articles published during the time frame of the survey. A greater percentage of the women faculty had not published an original article: 58% versus 47% for those with M.D. degrees, 43% versus 29% for those with M.D./Ph.D.s, and 29% versus 18% for those with Ph.D. degrees. Of those faculty who had published at least one original article, a greater percentage of the men than the women reported five or more original publications.

A researcher was defined in a previous publication[7] as someone who spent at least 20% of his or her time in research, had at least one original publication within a two-year period of time, and had external funds to support the research or assigned space in which to do the research; it was recognized that, at times, the researcher might be without funds or assigned space. The faculty in the present study who were the principal investigators on NIH grants best fulfilled these criteria. Similar results were found in this analysis when these data were considered by the gender of the researcher.

Summary and Discussion

The authors found striking differences, as well as some similarities, between the men and the women faculty members of departments of internal medicine in analyzing the data obtained from the research activity forms. The women as a group tended to be younger and lower in rank. More than half the reporting women faculty were less than 40 years old and more than two-thirds held the junior rank of instructor or assistant professor. In contrast, the majority of the men were between the ages of 40 and 59 and held the rank of associate professor or full professor. Notably, the percentage of women instructors and assistant professors exceeded the percentage of associate and full professors even within the 40–50-year age range. (A larger percentage of the women—18%—than of the men—7%—held the Ph.D. degree.)

The women with M.D. degrees reported significantly less NIH research support and more research support from non-NIH external funding sources than did the men with M.D. degrees. Even when the more mature age group (40–59) was examined, the men were significantly more likely than the women to have NIH funding. Furthermore, a greater percentage of the men faculty reported assigned research space, and the women faculty tended to have fewer publications than did the men during the two years surveyed.

Both the men and the women respondents tended to best fulfill the criteria for active researchers when they were principal investigators on NIH grants. Another clear similarity was that a large percentage of the faculty with M.D. degrees, both men and women, did not participate in a significant research effort: the majority of both groups reported spending less than 20% of their professional time in research activities, and almost one-third of both groups reported no research effort. Interestingly, a greater percentage of those with M.D./Ph.D. degrees than those with Ph.D. degrees reported no research effort.

In reflecting on these findings, the authors note that the number of students entering medical schools began to increase sharply in 1955, and the advent of Medicare and Medicaid thrust appreciable direct patient care responsibilities upon departments of medicine. Faculty recruitment has expanded since that time in order to meet the challenges of increased class size and increased numbers of patients. The findings of this study with regard to research efforts suggest that both women and men are being recruited to departments of medicine in appreciable numbers to fulfill the roles of patient care and teaching. However, this does not explain the discrepancies between the activities of the men and women faculty in the areas generally defined as "scholarly pursuits"—that is, those areas evidenced by publications and NIH grant support—at times of promotion and conferral of tenure.

Fletcher[9] cautioned in 1982 that academic women who are physicians should guard lest other activities suffocate their professional growth. The women faculty of internal medicine discussed in this study tended to be younger in age and lower in rank than were the men who were their counterparts; both these characteristics make the women particularly vulnerable in their academic careers. This study has not investigated the professional or personal activities that may have accounted for the difference between the ranks of the men and women in the 40–59-year age group. Further investigations are warranted to monitor as well as assess the progress of women in their academic careers.

The authors thank Linda Hough for data analyses, and Dr. David Babbott, Ms. Elizabeth Higgins, and Dr. Thomas Dial for their contributions to the original survey report. This work was supported by a grant from the Richard King Mellon Foundation.

References

1. Lister, J. The Role of Women in Medicine. *N. Engl. J. Med.* **300**(1979):1260–1261.
2. Koshland, D. E., Jr. Women in Science. *Science* **239**(1988):1473.
3. Relman, A. S. Here Come the Women. *N. Engl. J. Med.* **302**(1980):1252–1253.
4. Braslow, J. B., and Heins, M. Women in Medical Education. *N. Engl. J. Med.* **304**(1981):1129–1135.
5. Bickel, J. Women in Medical Education: A Status Report. *N. Engl. J. Med.* **391**(1988):1579–1584.
6. Beaty, H. N., Babbott, D., Higgins, E. J., Jolly, P., and Levey, G. S.: Research Activities of Faculty in Academic Departments of Medicine. *Ann. Intern. Med.* **104**(1986): 90–97.
7. Gentile, N. O., Jolly, P., Levey, G. S., and Dial, T. H. *Research Activity of Full-Time Faculty in Departments of Medicine.* Washington, D.C.: Association of Professors of Medicine and Association of American Medical Colleges, 1987, pp 1–56.
8. Mulvey, Holly, ed. *In the Marketplace: Work Patterns, Practice Characteristics and Incomes of Women Physicians.* American Medical Association Women in Medicine Project. Chicago, Illinois: American Medical Association, 1985.
9. Fletcher, S. W. Women Physicians: Old Times or a New Era? *The Pharos* **45** (Winter 1982):2–9.

A3273-8

MARCH 1990

ACADEMIC MEDICINE

Medical School Admission

Overview and Perspectives

Demographic Variables

Qualitative Variables

Traditional Predictors

Predicting Minorities' Success

The Interview

Moral Dilemmas

Summary and a Proposal

AAMC Data Report

VOLUME 65 • NUMBER THREE

MARCH 1990
VOLUME 65
NUMBER THREE

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

**Theme Issue:
Medical School Admission**

Foreword, *by Addeane S. Caelleigh* 133

Preface, *by Janine C. Edwards, Ph.D.* 134

Overview, *by George E. Miller, M.D.* 135

**Special Articles:
Medical School Admission**

Perspectives on Medical School Admission, *by William C. McGaghie, Ph.D.* 136

Demographic Variables in Medical School Admission, *by George Nowacek, Ph.D., and Larry Sachs, Ph.D.* 140

Qualitative Variables in Medical School Admission, *by William C. McGaghie, Ph.D.* 145

Traditional Predictors of Performance in Medical School, *by Karen J. Mitchell, Ph.D.* 149
Counterpoint: The Mismeasure of Medical Education, *by Norman D. Anderson, M.D.* 159

Predicting Minority Students' Success in Medical School, *by William E. Sedlacek, Ph.D., and Dario O. Prieto, M.Ed.* 161

The Interview in the Admission Process, *by Janine C. Edwards, Ph.D., Eugene K. Johnson, M.S., and John B. Molidor, Ph.D.* 167
Counterpoint: The Interview: One More Life?, *by Thomas C. Taylor* 177

Moral Dilemmas in Medical School Admission, *by Donnie J. Self, Ph.D.* 179

Help for the Gatekeepers: Comment and Summation on the Admission Process, *by Charles E. Spooner, Jr., Ph.D.* 183

A *Proposal*: A Matching Program for Medical School Admission, *by Martin A. Pops, M.D.* 188

Ideas for Medical Education — A Recruitment and Preparation Program for Hispanic Applicants, *by Louis S. Binder, M.D., Joseph Brown III, M.D., and Larry P. Jones, Ph.D.* 190

Book Reviews — *The Tutorial Process*, 191; *The Physician as Teacher*, 192; *The Physician Executive*, 192; *The Second Medical Revolution: From Biomedicine to Infomedicine*, 193; *The Financing of Biomedical Research*, 194; *Intimate Adversaries: Cultural Conflict Between Doctors & Women Patients*, 195; *A Model of Its Kind: A Centennial and Pictorial History of Johns Hopkins*, 196; *Making Medical Doctors: Science and Medicine at Vanderbilt since Flexner*, 198; *Yankee Surgeon: The Life and Times of Usher Parsons*, 200; *Handbook of Medical Library Practice, Vol. III*, 201

Letters to the Editor 202

From the President 203

National Policy Perspectives 204

AAMC Focus 206

Research Reports

ARTICLES

The Controllable Lifestyle Factor and Students' Attitudes about Specialty Selection, *by Richard W. Schwartz, M.D., John V. Haley, Ph.D., Cheryl Williams, Roy K. Jarecky, Ed.D., William E. Strodel, M.D., Byron Young, M.D., and Ward O. Griffen, Jr., M.D., Ph.D.* 207

The Role of Partners in Selecting a Residency, *by Robert M. Arnold, M.D., Carol Landau, Ph.D., Julia C. Nissen, Pharm.D., Steven Wartman, M.D., Ph.D., and Seth Michelson, Ph.D.* 211

SHORT PAPER

Personality Types of Family Practice Residents in the 1980s, *by Anita D. Taylor, M.A. Ed., Craig Clark, M.D., and Ann E. Sinclair, M.S.* 216

AAMC Data Report — The Declining Applicant Pool and Academic Quality, *by DeWitt S. Clark, Ph.D., Charles D. Killian, M.A., and Karen J. Mitchell, Ph.D.* 219

Bibliography

The February issue was the last in which *Academic Medicine* printed the monthly bibliography from the National Library of Medicine. Those readers who wish to access the library's MEDLINE database should call for information at 1 (800) 638-8480.

Publishing Services: **Hanley & Belfus, Inc.**
210 South 13th Street, Philadelphia, PA 19107

MARCH 1990
VOLUME 65
NUMBER THREE

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editors
Donald G. Kassebaum, M.D.
Edward J. Stemmler, M.D.
Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
Arizona Health Sciences Center

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
Georgetown University Hospital
Philip Anderson, M.D.
University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
Harvard Medical School
Eta Berner, Ed.D.
University of Alabama School of Medicine
Ruth Bulger, Ph.D.
National Academy of Sciences
Vincent A. Fulginiti, M.D.
Tulane University School of Medicine
Nancy Gary, M.D.
Albany Medical College
Beatrix Hamburg, M.D.
Mt. Sinai Hospital
Murray M. Kappelman
The University of Maryland School of Medicine
Ralph W. Muller
University of Chicago Hospitals
Herbert Pardes, M.D.
College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
University of Alabama School of Medicine
W. Vickery Stoughton
Toronto General Hospital
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center School of Medicine

Chairman-Elect
David H. Cohen, Ph.D.
Northwestern University

Immediate Past Chairman
John W. Colloton, M.A.
University of Iowa Hospitals and Clinics

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Ernst R. Jaffé, M.D.
Albert Einstein College of Medicine

Chairman, Council of Teaching Hospitals
Gary Gambuti
St. Luke's-Roosevelt Hospital Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, N.W., Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.

Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

RESEARCH REPORTS

ARTICLES

The Controllable Lifestyle Factor and Students' Attitudes about Specialty Selection

RICHARD W. SCHWARTZ, M.D., JOHN V. HALEY, Ph.D.,
CHERYL WILLIAMS, ROY K. JARECKY, Ed.D., WILLIAM E. STRODEL, M.D.,
BYRON YOUNG, M.D., and WARD O. GRIFFEN, Jr., M.D., Ph.D.

Abstract—Questionnaires were distributed to 346 fourth-year students in nine medical schools. The students were asked to state their selected specialty and to rank the importance that each of 25 influences, listed as questionnaire items, had had in making their choice of specialty. Factor analysis showed that particular items were significantly associated with particular factors. The first factor emphasized perceived lifestyle (items in this category gave importance to remuneration, personal time, and prestige); the second factor emphasized cerebral activities and a practice orientation; and the third factor stressed altruistic values and attitudes. The authors classified the selected specialties into three groups: those characterized as having a non-controllable lifestyle (NCL), those with a controllable lifestyle (CL), and surgery. (CL specialties were defined as those that allow the physician to control the number of hours devoted to practicing the specialty.) Data were analyzed using factor analysis, and analysis of variance, and the Scheffé method. Analysis indicated that the perceived lifestyle factor was most closely associated with the responses of those students choosing CL specialties. Furthermore, this factor received the highest total loading of the three factors from all the students, thus indicating the level of interest in lifestyle factors. Responses to items that defined the cerebral and practice factor were highest from the group of students choosing CL specialties and lowest from the group choosing NCL specialties. The NCL students scored highest in the altruism factor and the CL students scored the lowest. The surgery and NCL groups were similar in attitude patterns, and both were substantially different in attitude patterns from those of the CL groups. These findings were similar for the students from all institutions, regardless of variables such as geography or type of funding, and have implications for identifying students who may be most likely to enter primary care specialties. The results possibly reflect an increase in competitiveness among academically capable students for the limited number of first-year and second-year CL residency positions. Finally, results may reveal the social attitudes and values with which many present-day students tend to identify as compared with those of medical students who were graduated in the previous two decades. *Acad. Med.* 65(1990):207–210.

Traditionally, factors such as specialty content, following a role model, specialty prestige, opportunities for cognitive performance, personality characteristics, and future financial remuneration have been important in the choice of specialty training by graduating medical students. With changes that have occurred in American cultural values and concurrently in the practice of medicine, factors such as the control of work hours, the likelihood of litigation, length and mode of training, and employment opportunities have become increasingly important.

Recent trends in the specialty choice of medical students have indicated that primary care specialties, such as internal medicine and family practice, are no longer popular.[1,2] In another study,[3] some of the authors of the present study demonstrated that the most academically successful medical students (that is, those in the top 10% of their classes) are increasingly choosing "controllable lifestyle" (CL) specialties. (These are specialties that allow the physician to control the number of hours devoted to practicing the specialty.) Furthermore, in a second study[4] the authors have shown that young physicians engaged in primary care and surgical specialties are switching to CL-type specialties, thus increasing an already critical need for more primary care practitioners. Medical educators are searching for information that will enable medical school admissions committees to select a higher percentage of students who are more likely to choose a primary care specialty.

The purpose of this study was to disclose the relationship between (1) the kinds of influences that selected students said were important in their

---

Dr. Schwartz is assistant professor, Division of General Surgery, Department of Surgery; Dr. Haley is professor of behavioral science, Department of Behavioral Science; Ms. Williams is a medical student, class of 1990; Dr. Jarecky is professor of behavioral science and a special assistant to the dean; Dr. Strodel is professor of surgery and chief, Division of General Surgery; Dr. Young is professor of surgery and chairman, Department of Surgery; all are at the University of Kentucky College of Medicine, Lexington, Kentucky. Dr. Griffen is professor of surgery, Temple University, Philadelphia, Pennsylvania, and the executive director of The American Board of Surgery.

Address correspondence to Dr. Schwartz, Department of Surgery, University of Kentucky Chandler Medical Center, 800 Rose Street, Lexington, KY 40536-0084. Reprints will not be available.

specialty selections, and (2) the specialties that they selected.

Method

A 25-item questionnaire was developed that asked fourth-year students to list their specialty choice and to rate the degree of importance to that choice of the following: (1) ten influences, such as clerkship experiences, classmates, and clinical role models; (2) ten standard specialty characteristics, such as types of patients to be cared for, number of hours per week of work, and presumed level of specialty prestige; and (3) five personal long-term goals, such as to provide community service, to enter academic medicine, and to practice in a rural area. The questionnaire was sent in 1988 to nine U.S. medical schools* and distributed to their fourth-year students, a total of 1,082 students. Four schools were located in the Midwest, two in the South, two in the Southwest, and one in the Northeast. The questionnaires were completed before graduation by 346 of the students, or 32% of the total group. The proportions of fourth-year students who responded from the individual schools varied widely; for example, 48% responded from the University of Arkansas and 44% from the University of Arizona, but only 7% returned completed questionnaires from the University of Louisville. The other schools had percentages of return ranging from 24% to 38%. For some of the schools, the number of respondents might be considered representative of the fourth-year class, but for others this was not the case. Therefore, no claims can be made that the respondents were representative of the medical schools that cooperated in the study.

The respondents rated each item on a Likert-type scale of 1 to 5, using the following labels: 1 = of no importance; 2 = of minor importance; 3 = of some importance; 4 = very important; and 5 = critically important. They also indicated their specialty choice. Eight students who were undecided were eliminated from the analysis.

The specialties selected were subdivided by the authors into three groups: non-controllable lifestyle (NCL); controllable lifestyle (CL); and surgery. Surgery, although logically a part of the NCL group, was considered separately because of a special interest in this specialty on the part of the investigators. The NCL specialties were internal medicine, family practice, pediatrics, and obstetrics-gynecology. The CL specialties were anesthesiology, dermatology, emergency medicine, neurology, ophthalmology, otolaryngology, pathology, psychiatry, and radiology. "Surgery" comprised general surgery and the surgical specialties.

The 25 questionnaire items were correlated and factor-analyzed using the principal-components method. The factors were then rotated to simple structure using the Varimax procedure. Scores were obtained for each student on each factor by summing the student's scores on the questionnaire items that defined that factor. These scores were then compared across specialty groups using a one-way analysis of variance. Additional comparisons were made across medical schools to see whether school bias accounted for any of the differences found.

Results

Three factors emerged. Factor I had six questionnaire items with high loadings. (A high loading reflects the magnitude of the correlation of that item with the factor. In this study, a correlation of .30 or greater was considered high.) Since all of these items related to remuneration, personal time, and prestige, this factor was defined as the "perceived lifestyle" factor. Four items had high loadings for factor II, which was defined as the "cerebral and practice orientation" factor because the items emphasized opportunities for cerebral activities and for a variety of practice experiences. Factor III had four items with high loadings also and was defined as the "altruism" factor to reflect the items' emphasis on altruistic values and attitudes. Table 1 lists the items that defined each factor and the corresponding factor loadings.

Each student was given a score for each factor. If a student did not respond to all the items defining a factor, none of that student's responses were used in the analysis of that factor. This was the case for 12 students for the factor I analysis, 82 for factor II, and 30 students for the factor III analysis.

An analysis of variance was used to compare the scores of the three specialty groups for each of the three factors. (Hereafter, these scores are called "factor scores.") Using the Scheffé method, a posteriori comparisons were made. Table 2 presents the means and standard deviations for each of the scores. The CL group placed greater emphasis on the items that defined the perceived lifestyle factor than did the NCL and surgery groups. The latter two groups placed about the same emphasis on these items. The cerebral and practice orientation items were rated the highest by the CL group, followed by the surgery group, with the NCL group scoring the lowest. Significant differences regarding these items were found only between the CL and the NCL groups. Finally, the NCL group gave the highest ratings to the altruism items and the CL group gave the lowest, with the surgery group in the middle. Both the surgery and the NCL groups differed significantly from the CL group regarding these items, although they did not differ significantly from each other.

---

*Washington University School of Medicine, St. Louis, Missouri; East Tennessee State University James H. Quillen College of Medicine, Johnson City; the University of Arizona College of Medicine, Tucson; the Medical College of Georgia School of Medicine, Augusta; State University of New York Health Science Center at Syracuse College of Medicine, Syracuse; Bowman Gray School of Medicine of Wake Forest University, Winston-Salem, North Carolina; the University of Louisville School of Medicine, Louisville, Kentucky; the University of Arkansas College of Medicine, Little Rock; and the University of Kentucky College of Medicine, Louisville.

The factor scores of the groups of students at the nine medical schools were compared by using a one-way analysis of variance. There were a significant differences among the schools for factors II and III but not for factor I. The purpose of this comparison was to see whether a school bias could account for differences found among the CL, NCL, and surgery groups. A comparison of the three groups within each school indicated this was not the case. For factors I and III, all nine schools followed the same pattern observed when the scores of the schools were pooled together. Seven of the nine schools followed the same pattern for factor II. These patterns were not always significant, due to loss of power resulting from the reduced sample size. Nevertheless, they were consistent with the main findings, suggesting that the differences observed in the pooled comparisons were true for the students who filled out the questionnaire from each of the nine medical schools.

Table 1

Questionnaire Items Defining Each of Three Factors and the Degree of Loading of Each Item for Its Associated Factor, 1988*

Factors and Associated Items	Factor Loading†
Perceived lifestyle factor	
Potentially high remuneration	.54
Number of work hours/week in practice	.84
Adequate time for pursuit of avocational activities during practice	.79
Perceived number of call nights during practice	.80
Presumed level of specialty prestige	.48
Length of residency program	.45
Cerebral and practice activities factor	
Clerkship experience at the respective medical centers	.50
Clinical faculty role models	.51
To enter academic medicine	.58
To practice in a large urban area	.47
Altruism factor	
Non-faculty physician role models	.45
Types of patients to be cared for	.50
To provide community service	.48
To practice in a rural area	.56

*A total of 346 fourth-year medical students from nine medical schools were given a 25-item questionnaire that asked them to rate various influences basic to selecting their specialty. The items were correlated and factor-analyzed. Three factors emerged: the "perceived lifestyle" factor (items in this category gave importance to remuneration, personal time, and prestige); the "cerebral and practice" factor, defined by items that emphasized cerebral activities and a practice orientation; and the "altruism" factor, which stressed altruistic values and attitudes.

†An item's factor loading reflects the magnitude of the correlation of that item with the factor. In this study, a correlation of .30 or greater was considered high.

Table 2

Means of Students' Scores on Groups of Questionnaire Items, Each of Which Defined One of the Three Factors Describing the Students' Orientation for Each of the Three General Types of Specialties Chosen, 1988*

	Factor								
	Perceived Lifestyle‖			Cerebral and Experiential¶			Altruism**		
General Specialty Type	Mean	SD	No. Students Choosing Type of Specialty	Mean	SD	No. Students Choosing Type of Specialty	Mean	SD	No. Students Choosing Type of Specialty
Controllably lifestyle†	20.74	5.00	113	11.44	3.11	87	11.23	2.62	108
Noncontrollable lifestyle‡	18.42	4.82	165	9.26	4.24	129	13.59	2.73	152
Surgery§	18.21	4.15	48	10.68	3.38	40	12.94	2.61	48

*A total of 346 fourth-year medical students from nine medical schools were given a 25-item questionnaire that asked them to rate various influences basic to selecting their specialty. The items were correlated and factor-analyzed. Three factors emerged: the "perceived lifestyle" factor (emphasizing remuneration, personal time, and prestige); the "cerebral and practice" factor (emphasizing cerebral activities and practice orientation); and the "altruism factor" (emphasizing altruistic values and attitudes). Each student was given a score for each factor; each score was the sum of the ratings on the items with high loadings for the particular factor. Each mean score for each factor listed in the table is the mean of the scores of the specific group of students whose specialties were of a particular general specialty type.

†The following specialties were classified by the authors as "controllable lifestyle" (CL) specialties (that is, those that allow the physician to control the number of hours devoted to practicing the specialty): anesthesiology, dermatology, emergency medicine, neurology, ophthalmology, otolaryngology, pathology, psychiatry, and radiology.

‡The following specialties were classified by the authors as "noncontrollable lifestyle" (NCL) specialties: internal medicine, family practice, pediatrics, and obstetrics-gynecology.

§Surgery, although a part of the noncontrollable lifestyle group, was considered separately because of a special interest in this specialty on the part of the authors. The surgery group comprised general surgery and the surgical specialties.

‖For the perceived lifestyle factor, CL versus NCL specialties was significant ($p < .05$), and CL versus surgery specialties was significant ($p < .05$).

¶For the cerebral and practice factor, CL was significantly greater than NCL ($p < .05$).

**For the altruism factor, NCL and surgery were significantly greater than CL ($p < .05$).

Discussion

Definite trends have evolved in the investigation of the significant factors affecting medical student specialty choice.[5] In the 1960s and early 1970s, interest centered on cognitive factors. Student personality characteristics and attitudes as well as perceived specialty status dominated the late 1970s. Since the early 1980s, investigators have concentrated on examining factors that occur during or after medical school. Perhaps the major examples of recent investigation of this type are the studies reporting the obvious combination of medical student aversion towards primary care and attraction for CL specialties.[1-4]

Two recent studies have suggested that medical students are increasingly choosing CL specialties.[3,4] Since these CL specialties are becoming more competitive, top students are selecting them rather than the NCL specialties. As a result, the quality of students selecting primary care and surgery specialties will begin to decrease. This trend has already affected primary care, as evidenced by a decline in student interest since 1981.[1,2]

The results of the present study suggest that the students choosing a CL specialty placed more emphasis on perceived lifestyle factors than did the students choosing either primary care or surgery specialties. Concurrently, the CL group placed less emphasis on items pertaining to altruism than did the other two groups. This is in direct contrast to a 1971 study by Paiva and Haley,[6] in which students choosing internal medicine and surgery ranked social values lowest. The CL group placed the highest emphasis on the cerebral and practice orientation items. This relates favorably to work performed by Essex-Sorlie and Sorlie[5] that demonstrated that students selecting anesthesiology, pathology, and radiology recorded the highest mean performances on variables associated with prematriculation achievement. Their study also showed that students choosing internal medicine and surgery achieved the highest mean scores during medical school matriculation. This appears to be changing as top medical students are increasingly choosing CL specialties.[3]

One major finding of the present study was that there are significant differences in the importances of the factors examined. Students selecting CL specialties clearly rated themselves highest in items classified by the perceived lifestyle factor and lowest in those under the altruism factor. Zimet and Held[7] in 1975 performed a longitudinal study of medical student career choices and focused on beliefs about status, social attractiveness, and similarity to self of various specialty practitioners. The results, which were representative of other such studies, demonstrated that surgery ranked the highest in status while ranking the lowest in social attractiveness; not surprisingly, family practice ranked conversely to surgery on both scales. Even with those clear differences, the majority of students, regardless of future specialty choice, saw themselves as personally being most similar to the "typical" family practitioner, in terms of the characteristics they attributed to this type of practitioner. Such characteristics included warmth, concern for patients, and orientation to the other person rather than to a procedure or test. The data of the present study strongly imply that the current generation of medical students are either much more honest than their predecessors or much less altruistic. Given the current "yuppie" atmosphere prevalent in today's society, the authors suspect that, disturbingly, the latter hypothesis is true.

These results may be used to identify a group of medical students whose initial interest in medicine appears to lie outside the primary care fields. Other authors[6,7] have suggested that appropriate examination of students' attitudes about both themselves and medical specialties may differentiate, upon medical school entrance, the eventual specialty choices of medical students. This information, combined with cognitive performance at both the undergraduate and medical school levels, is needed to recruit and admit medical students whose eventual career intention is that of primary care. Faculty intervention tactics during medical school matriculation, such as appropriate inpatient and outpatient clinical experiences and role modeling, may be needed to make specialties such as family practice, internal medicine, pediatrics, and surgery more attractive to the current generation of medical students.

The authors thank John C. Herweg, M.D., Washington University School of Medicine; Randolph Manning, Ed.D., then of the East Tennessee State University James H. Quillen College of Medicine; Andrew M. Goldner, Ph.D., University of Arizona College of Medicine; Terrence T. Kuske, M.D., Medical College of Georgia; A. Geno Andreatta, State University of New York Health Science Center at Syracuse College of Medicine; Nat E. Smith, M.D., Bowman Gray School of Medicine; Hiram Polk, M.D., University of Louisville School of Medicine, and Ture W. Schoultz, then of the University of Arkansas College of Medicine, for their unstinting assistance in securing the data basic to this study.

References

1. *NRMP Directory: Including Hospitals and Programs Participating in the Matching Program for 1987.* Evanston, Illinois: National Resident Matching Program, 1987, p. 131.
2. Jones, Robert F. American Medical Education: Institutions, Programs, and Issues. AAMC Staff Report. Washington, D.C.: Association of American Medical Colleges, 1989, p. 21, Figure 12.
3. Schwartz, R. W., Jarecky, R. K., Strodel, W. E., Haley, J. V., Young, A. B., and Griffen, W. O., Jr. Controllable Lifestyle: A New Factor in Career Choice by Medical Students. *Acad. Med.* **64**(1989):606-609.
4. Schwartz, R. W., et al. Career Change: In Quest of a Controllable Lifestyle. *J. Surg. Res.* **47**(1989):189-192.
5. Essex-Sorlie, D., and Sorlie, W. The Relationship between Cognitive Factors and Medical Student/Residency Selection. In *Research in Medical Education: 1984. Proceedings of the Twenty-Third Annual Conference*, S. Kerby, compiler, pp. 173-178. Washington, D.C.: Association of American Medical Colleges, 1984.
6. Paiva, R. E. A., and Haley, H. B. Intellectual, Personality, and Environmental Factors in Career Specialty Preferences. *J. Med. Educ.* **46**(1971):281-288.
7. Zimet, C. N., and Held, M. L. The Development of Views of Specialties During Four Years of Medical School. *J. Med. Educ.* **50**(1975):157-166.

APRIL 1990
VOLUME 65
NUMBER FOUR

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Invited Articles

The Education of Medical Students: Sounds, Alarums, and Excursions, *by Daniel D. Federman*, M.D. 221

An Evolving Agenda for the AAMC's Second Century, *by Louis Sullivan*, M.D. 226

Restructuring Education and Its Impact on Medical Education, *by Lauro F. Cavazos*, Ph.D. 230

Ideas in Medical Education

Using Classic Mystery Stories in Teaching, *by Stephen H. Sheldon*, D.O., *and Peter A. Noronha*, M.D. 234

Academic Computing in Medicine

Status Report on the NBME's Computer-Based Testing, *by Stephen G. Clyman*, M.D., M.S., *and Nancy A. Orr*, Ph.D. 235

From the President 242

AAMC Focus 243

National Policy Perspectives 244

Commentary

A Plea for Positive Thinking about Modern American Medicine, *by Ernst Jaffé*, M.D. 246

Research Reports
ARTICLES

Investigating the Fairness of the National Residency Matching Program, *by Yufei Yuan*, Ph.D., *and Amiram Gafni*, D.Sc. 247

The Reliability of the Proposal Review Process for a Summer Research Fellowship Program, *by Antoinette S. Peters*, Ph.D., *and Thomas D. Flanagan*, Ph.D. 254

Variation in Lengths of Training among Residents Enrolled in Pediatrics and Internal Medicine Training Programs, *by Robert A. Hoekelman*, M.D., *and Ruth M. Parker*, M.D. 257

Surveying Graduates of Combined Internal Medicine–Pediatrics Residency Programs, *by Frank* M. *Biro*, M.D., *Matthew* W. *Gillman*, M.D., *Ruth* M. *Parker*, M.D., *Philip* R. *Khoury*, M.S., *and David* M. *Siegel*, M.D. 266

SHORT PAPERS

Mentors in Graduate Medical Education at the Medical College of Wisconsin, *by Robert* A. *Kirsling*, M.S., *and Mahendr* S. *Kochar*, M.D., M.S. 272

A New Method of Assessing the Interpersonal Skills of Surgeons, *by Kenneth* W. *Burchard*, M.D., *and Pamela* A. *Rowland-Morin*, Ph.D. 274

AAMC Data Report

Women on Faculties of U.S. Medical Schools, 1978–1989, *by Brooke* E. *Whiting*, Ph.D., *and Janet Bickel*, M.A. 277

Publishing Services: Hanley & Belfus, Inc.
210 South 13th Street, Philadelphia, PA 19107

APRIL 1990
VOLUME 65
NUMBER FOUR

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
Edward J. Stemmler, M.D.
Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
Arizona Health Sciences Center

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
Georgetown University Hospital
Philip Anderson, M.D.
University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
Harvard Medical School
Eta Berner, Ed.D.
University of Alabama School of Medicine
Ruth Bulger, Ph.D.
National Academy of Sciences
Vincent A. Fulginiti, M.D.
Tulane University School of Medicine
Nancy Gary, M.D.
Albany Medical College
Beatrix Hamburg, M.D.
Mt. Sinai Hospital
Murray M. Kappelman, M.D.
The University of Maryland School of Medicine
Ralph W. Muller
University of Chicago Hospitals
Herbert Pardes, M.D.
College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
University of Alabama School of Medicine
W. Vickery Stoughton
Toronto General Hospital
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
David H. Cohen, Ph.D.
Northwestern University

Chairman-Elect
William T. Butler, M.D.
Baylor College of Medicine

Immediate Past Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center School of Medicine

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Joe Dan Coulter, Ph.D.
University of Iowa College of Medicine

Chairman, Council of Teaching Hospitals
Raymond G. Schultze, M.D.
University of California, Los Angeles Medical Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.

Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036; (202) 828-0416.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

RESEARCH REPORTS

ARTICLES

Investigating the Fairness of the National Resident Matching Program

YUFEI YUAN, Ph.D., and AMIRAM GAFNI, D.Sc.

Abstract—The National Resident Matching Program (NRMP) was established in the early 1950s to bring order and fairness to a previously chaotic application process for internship and residency positions. Over the years many reservations were raised about the fairness of the process, specifically, that hospital programs are doing better than students are (i.e., programs obtain preferred residents more often than students receive preferred programs). This paper presents an analysis of the results of the 1986 Match. The findings do not support the claim that the results of the Match are biased in a way that favors programs. Overall, the students' success was greater (i.e., receiving on average higher-ranked choices) than was the hospital programs' success. In fact, in 20 of 22 specialty programs, the students' degree of success was greater than or equal to the success of the programs. This finding raises the question of fairness toward programs rather than students. The authors analyze factors that affected both the hospitals' and the students' degrees of success in the 1986 Match and suggest strategies for improving the Match results. *Acad. Med.* 65(1990):247–253.

The National Resident Matching Program (NRMP) was established in the early 1950s to bring order and fairness to a previously chaotic application process for internship and residency positions.[1,2] In brief, from the turn of the century until 1945, the competition by hospitals for interns and residents manifested itself in a race to sign employment contracts earlier and earlier in medical students' careers. In some cases students were subjected to direct pressure from programs and were forced to decide prematurely. The adoption of a centralized matching mechanism seemed to resolve most of these problems.

The NRMP requires each of the participants (students and hospital program directors) to supply a rank-order list of preferences. Based on these rank-order lists, a computerized algorithm is used to assign the students to hospital positions. The assignment algorithm used is described in general terms in the NRMP directory[3] and is equivalent to the one described by Gale and Shapley.[4] The use of the algorithm provides a "stable solution," namely, the outcome of the matching process is such that after the Match, there exists no student and hospital program (who are not assigned to each other) who would *both* prefer each other to their assigned partners.

In spite of this theoretically stable solution, many reservations were raised during the years about the fairness of the NRMP process. For example, some said that residency program directors understand the Match system better (stemming from past experience resulting from participating in many Matches) and thus are able to receive better students because they know which ones to rank higher[5]; that due to lack of satisfaction with the "official" program (NRMP), "unofficial" arrangements have appeared[6]; that there is dishonesty in some recruitment interviews from program directors[7]; and that the current computer algorithm treats students and programs unequally in a way that favors programs.[8]

Nevertheless, the NRMP has been praised as an equitable system of residency matching. Some proponents insist that it is a far fairer system than the routine admission-rejection system used by colleges, medical schools, or other schools.[9] Since most complaints indicate a bias in favor of hospital programs, it is important to examine whether the criticism is supported by evidence. In this paper we present an analysis of the results of the 1986 NRMP.

Method

The data about the hospital programs and students who participated in the 1986 Match were provided to us by the NRMP with special permission. Unless otherwise stated, "student" is being used as a general term for all the types of applicants: U.S. senior students, U.S. citizens who are graduates of foreign medical schools, and others. In order to keep the data confidential, the students' and hospitals' names were omitted. In addition, the students' and programs' identifications were reformulated so they could

---

Dr. Yuan is assistant professor, Area of Management Science and Information Systems, Faculty of Business, and Dr. Gafni is associate professor, Centre for Health Economics and Policy Analysis, Department of Clinical Epidemiology and Biostatistics, Faculty of Health Sciences; both at McMaster University, Hamilton, Ontario, Canada.

Correspondence and requests for reprints should be sent to Dr. Gafni, Department of Clinical Epidemiology and Biostatistics, Room 2C12A, McMaster University, Health Sciences Centre, 1200 Main Street West, Hamilton, Ontario, Canada L8N 3Z5.

not be recognized.

The data were organized in three files: the students' ranking file, the hospital programs' ranking file, and the Match result file. The students' ranking file provided ranking information about all the students who participated in the Match. Each record consisted of the student's identification, programs listed by the student, the rank assigned to the student, whether the student was a U.S. senior, whether he or she was part of a couple for matching, and whether the student was active in the Match (that is, did not withdraw from the Match process). Similar information for the hospital programs is found on the hospital programs' ranking file. Each record consisted of the hospital program's identifications, the type of program (specialty), the students listed by the program, the ranks assigned to them, the number of positions offered, and the number of positions filled. The Match result file provided information about all the Match pairs (the programs and the assigned students), the unmatched students, and the unfilled positions.

The analysis consisted of three tasks: (1) To compare the students' and the hospitals' Match successes; (2) to identify factors that affected the hospitals' Match success; and (3) to identify factors that affected the students' Match success.

To compare the students' success with the hospitals' success we calculated the mean value and standard deviation of the program's rank for the matched student and the student's rank for the matched position. Hospital programs have different quotas. Thus, for a student to be ranked as the third candidate by a hospital program that offers only one position is not the same as being ranked third candidate by a hospital program that offers ten positions. Therefore, to make meaningful comparisons, we standardized the rank order by using the following strategy: the rankings submitted by a hospital program were divided into subgroups. The size of each subgroup was equal to the number of positions (quota) offered by that program. Thus, the first subgroup consisted of a number of students — equal to the quota of that hospital program — who were ranked highest on its list. This subgroup was considered the "standardized first choice." Similarly, the students in the second-highest ranking subgroup were considered the "standardized second choice," and so on. For example, if a program's quota were five, the first- to the fifth-ranked students listed by the program would be considered the "standardized first choice" students for its five positions. The sixth- to the tenth-ranked students would be considered the "standardized second choice" students, and so on. Notice that by using this method, we assumed that (1) all the positions offered by a hospital program were identical, and (2) all candidates in each subgroup were equally favorable for all the positions of the hospital program. In other words, if the hospital filled some positions with candidates from the first-choice subgroup, this would be the same as filling one of these positions with a first-choice candidate, and so on.

We calculated the difference between the student's rank and the hospital program's rank (standardized) for each matched pair for the system as a whole (all specialties together) and for each specialty. We employed the *t*-test to find out whether the mean of the rank difference was zero for all specialties as a whole and for each specialty. Finally, we tried to ascertain whether the mean rank differences were affected by the size of the market and/or demand-supply relations for each program specialty. This was done by calculating the corresponding correlation coefficients.

The Match results are different for individual specialties. To identify the factors that affected individual programs' Match successes, we had to analyze the Match results within each specialty. To keep our study manageable, three program specialties were chosen: internal medicine, surgery, and obstetrics-gynecology. These specialties were selected because a large number of positions were offered to students in each of them. Also, as shown later, they represented three different cases: one where the students' success in general was better than the hospital programs' success (internal medicine), one where there was no statistical difference between the students' and the programs' successes (obstetrics-gynecology), and one where programs' success in general was better than students' success (surgery).

In our analysis we used the filling rate and mean of each program's ranking to measure each program's success. The filling rate was defined as the ratio of the number of positions filled to the number of positions offered by a program. The mean of each program's ranking was the average of all the standardized program rank orders assigned to the students who were matched to that program. We concentrated mainly on those factors that were easy to measure and thus could be used most easily in practice to improve future success. We selected three factors as independent variables (possible factors affecting success): (1) the number of candidates listed per position, (2) the number of positions offered, and (3) the mean rank order assigned to the program by students who applied for the program. Pearson correlation coefficients between each program's success measures and the independent variables were calculated and a general linear model and *F*-distribution tests were used to test how these factors affected each program's success.

To identify the variables that affected the students' success, we considered only those students who were U.S. seniors, who applied for only one type of program (specialty), and were not a member of a couple participating in the Match. This subgroup was relatively homogenous and contained the majority of the students (about 70% of all applicants and 85% of all matched applicants). For the reasons previously mentioned, the analysis was carried out only for internal medicine, surgery, and obstetrics-gynecology (about 55% of all positions offered by all specialty programs).

To indicate a student's success or failure to obtain a position, we used a dichotomous variable, where 1 indi-

Table 1

Measurements of Students' and Programs' Successes in the 1986 Match*

Program Specialty†	Measurements‡								
	Positions Filled	A/P Ratio	MPR Mean	MPR SD	MSR Mean	MSR SD	DIFF Mean	DIFF SD	*p*-value from *t*-Test
Specialties for which students' or programs' success is demonstrated									
Internal medicine	5,985	8.8	3.25	2.74	2.28	2.60	1.24	3.47	.0001
Anesthesiology	280	5.7	2.56	2.21	1.68	3.17	0.88	3.66	.0001
Pediatrics	1,723	9.0	2.95	2.33	2.22	2.05	0.74	2.72	.0001
Family practice	1,960	6.9	2.62	1.80	1.92	1.86	0.70	2.32	.0001
Radiology—diagnostic	390	8.4	2.79	2.08	2.31	1.94	0.47	2.49	.0002
Psychiatry	782	6.1	1.63	1.09	1.43	1.18	0.19	1.40	.0001
Orthopedic surgery	355	15.4	2.01	1.39	2.58	2.76	−0.57	2.76	.0001
Surgery	1,994	9.2	2.23	1.57	3.39	4.12	−1.15	4.06	.0001
Specialties for which students' or programs' success is not demonstrated									
Obstetrics-gynecology	974	11.2	2.76	2.02	2.95	3.71	−0.19	3.82	.1193
Transitional	1,022	7.0	2.44	0.25	2.34	0.16	0.10	0.26	.6577
Emergency medicine	280	11.2	2.18	1.18	2.23	2.05	−0.05	2.08	.6883
Pathology	276	5.7	1.82	1.23	1.78	2.49	0.04	2.58	.8154
Urology	77	1.7	1.04	1.82	1.03	2.67	0.01	2.90	.2854
Physical medicine and rehabilitation	74	8.2	2.00	1.48	2.39	2.87	−0.39	2.99	.2629
Otolaryngology	40	1.4	1.00	—	1.00	—	—	—	—
Neurology	29	1.6	1.10	0.31	1.07	0.37	0.03	0.33	.5728
Neurologic surgery	29	1.5	1.00	—	1.00	—	—	—	—
Ophthalmology	18	2.6	1.22	0.43	1.11	0.47	0.11	0.47	.3313
Radiology—therapeutic	14	4.4	1.29	0.61	1.21	0.58	0.07	0.27	.3356
Plastic surgery	6	16.8	1.17	0.41	1.50	0.84	−0.33	0.82	.3632
Preventive med/public health	5	8.6	1.60	0.89	2.20	1.30	−0.60	1.82	.5012
Dermatology	5	7.9	1.00	—	1.00	—	—	—	—
Overall	16,318	8.5	2.83	2.28	2.34	2.76	0.49	3.28	.0001

*The authors analyzed data provided to them by the National Resident Matching Program to determine whether the students or the hospitals' specialty programs had greater success in the 1986 Match and to identify the factors contributing to the success of each group. The means of the rankings presented in the table (in the MPR and MSR columns) indicate that for six of the eight specialty programs for which a group's success could be shown, the students had greater success than did the programs. (Note: higher ranks—meaning greater success—are shown by smaller numbers.)

†Data about the specialty programs from the 1986 Match are presented here in two groups: the top group comprises those specialty programs for which the mean difference between the program's rank and the student's rank for the matched pairs was found to be significantly different from zero; the bottom group comprises the specialty programs where this difference was not significantly different from zero. In the bottom group, there is no way to ascertain, from the available data, whether one group was more successful than the other for a given specialty program.

‡The column headings are defined as follows: positions filled = number of positions filled in each specialty program; A/P ratio = number of applications per number of positions available; MPR = the mean rank order assigned to matched students by the programs; MSR = the mean rank order assigned to programs by the students; DIFF = the mean difference between the program's rank and the student's rank for the matched pair.

cated a student who was matched to a position and 0 indicated one who was unmatched. The independent variables considered were the number of programs listed by the student and the highest standardized program rank assigned to the student by all programs that listed him or her. Pearson correlation coefficients between the three variables were calculated for the three specialties considered in this study. A general linear model and *F*-distribution tests were used to test the impact of those variables.

Results

In Table 1 we present and compare the measurements of the students' and the programs' successes in the 1986 Match for all the specialty programs considered in the Match process that year. The students' mean rank for the position to which he or she was assigned was 2.34 (SD = 2.76). The programs' mean rank for a student assigned to an available position was 2.83 (SD = 2.28). A *t*-test performed for pair comparison shows that the mean of the rank difference is not zero ($p \leq .0001$). Notice that higher ranks are represented by smaller numbers; thus a smaller mean rank indicates better performance.

For Table 1 the different specialty programs were divided into two groups. The first group includes all programs where the mean difference between the program's rank and the student's rank for the matched pairs was found to be significantly different from zero. The second group includes the specialty programs where this difference was not found to be significantly different from zero. Notice

Table 2

Pearson Correlation Coefficients between Hospitals' Success Variables and Hospitals' Choice Variables, 1986 Match*

Choice Variable†	Success Variables,† by Specialty Program: Internal Medicine FIL	Internal Medicine MPR	Surgery FIL	Surgery MPR	Obstetrics-Gynecology FIL	Obstetrics-Gynecology MPR
CPR	.36770	.59637	.40585	.26764	.37676	.33335
Significance level‡	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)	(.0001)
MSR	.06070	.18739	.02960	.21945	.02050	.10479
Significance level‡	(.1024)	(.0001)	(.5302)	(.0001)	(.7421)	(.1011)
QUOTA	.14821	−.00357	.10016	−.13264	.26564	−.22379
Significance level‡	(.0001)	(.9250)	(.0321)	(.00600)	(.0001)	(.0004)

*The authors analyzed data provided to them by the National Resident Matching Program to determine whether the students or the hospitals' specialty programs had greater success in the 1986 Match and to identify the factors contributing to the success of each group. The correlations of the data above show that in all three specialty programs, the filling rate (FIL) was affected most by the number of candidates listed per residency (CPR). Program size (QUOTA) seems to have been much less important. The students' evaluation of the program (MSR) does not seem to have affected the filling rate. (See the text for additional interpretations of these data and an evaluation of the explanatory power of the three variables.)

†The variables are defined as follows: FIL = filling rate; MPR = the mean rank order assigned to matched students by the programs; CPR = the number of candidates listed per position; MSR = the mean rank order assigned to programs by the students; QUOTA = the number of positions offered by a program.

‡The numbers in parentheses show the significance level for testing (F-test) such that the coefficient = 0.

that in the first group, the students' success was greater than the programs' in six of the eight specialties. In two specialties (orthopedic surgery and surgery) the programs' success was greater.

Table 1 also contains information about the numbers of positions filled; these numbers can serve as indicators of the size of the market in 1986. We also present data on the ratio of the number of applications to the number of positions available in each program specialty. Notice that since every student could apply to more than one position, the number of applications was greater than the number of applicants (overall, the ratio of active applicants to positions in the 1986 Match was 1.1). We tried to explore whether differences in market size and/or the ability of programs to attract candidates (represented by the ratio of the number of applications to positions) could explain the differences in the mean ranking. No statistically significant relations were found between those variables and mean difference in ranking between matched programs and students.

Table 2 presents the results of our analysis of the three different variables (CPR, QUOTA, and MSR—defined below) that affected the hospital programs' success. An important finding is that in all three specialty programs, the filling rate was affected most by the number of candidates listed per residency position (CPR). This might indicate either the "aspiration level" of accepting a resident or the demand for the program, or both. Program size (QUOTA) seems to have been much less important. The applying students' evaluation of the program, measured by the students' average ranking (MSR), does not seem to have affected the filling rate. It is important to mention, however, that, based on a factor analysis using a general linear model, the explanatory power of these three variables together is about 40% in each of the three specialty programs ($R^2 =$ 0.383, 0.399, 0.392 for internal medicine, surgery, and obstetrics-gynecology, respectively).

The effect of the three variables just mentioned on the hospital programs' satisfaction from the students they received (measured by the average ranks of the students matched—the MPR variable) is also presented in Table 2. Again, the number of students listed per position seems to have been an important variable, but with negative impact. However, it was much more important in the case of internal medicine, less so in the case of obstetrics-gynecology, and even less in the case of surgery. The students' evaluation of the program seems to have affected the programs' satisfaction with the students they received, while the size of the program (QUOTA) does not seem to have had any effect. This is different from the effects of these two variables on the filling rate. Using the same type of analysis, we found that, compared with the former analysis, the three variables presented in Table 2 represent only a smaller portion of the explained variance ($R^2 = 0.237$, 0.167, 0.251 for internal medicine, surgery, and obstetrics-gynecology, respectively).

The number of programs listed by the student was found to have been an important factor in determining his or her chance to find a position. It is interesting to note that this factor was more important for internal medicine ($R^2 = 0.371$) and obstetrics-gynecology ($R^2 = 0.350$) than for surgery ($R^2 = 0.183$). The evaluation of students by programs was also found to be an important factor, but less so than the variable just discussed in determining the chances of a student to find a position. This factor, like the one previously discussed, was more important for obstetrics-gynecology ($R^2 = -0.214$) and internal medicine ($R^2 = -0.194$) than for surgery ($R^2 =$ -0.135). The significance level for rejecting ($F =$ test) such that the coefficient equals zero was 0.0001 in all cases.

Table 3 further illustrates how the number of candidates per position af-

Table 3

Impact of the Number of Candidates Listed per Position (Offered by Internal Medicine Programs) on the Programs' Success in Having All Available Positions Filled by the 1986 Match*

	Number of Candidates Listed per Position					
	0-29	3-5.9	6-8.9	9-11.9	12 or More	
Programs' Filling Rate†	No., %'s of Programs‡	No., %'s of Programs‡	No., %'s of Programs‡	No., %'s of Programs‡	No., %'s of Programs‡	Total and % All Programs§
0%	20 62.5% 22.7%	9 28.1% 4.1%	2 6.3% 1.1%	1 3.1% 1.0%	0 0.0% 0.0%	32 4.4%
10%	3 42.9% 3.4%	3 42.9% 1.4%	1 14.3% 0.6%	0 0.0% 0.0%	0 0.0% 0.0%	7 1.0%
20%	3 50.0% 3.4%	3 50.0% 1.4%	0 0.0% 0.0%	0 0.0% 0.0%	0 0.0% 0.0%	6 0.8%
30%	15 42.9% 17.1%	16 45.7% 7.3%	1 2.9% 0.6%	3 8.6% 3.1%	0 0.0% 0.0%	35 4.8%
40%	10 52.6% 11.4%	7 36.8% 3.2%	2 10.5% 1.1%	0 0.0% 0.0%	0 0.0% 0.0%	19 2.6%
50%	4 14.3% 4.6%	16 57.1% 7.3%	7 25.0% 3.9%	0 0.0% 0.0%	1 3.6% 0.7%	28 3.8%
60%	2 8.3% 2.3%	15 62.5% 6.3%	6 25.0% 3.4%	0 0.0% 0.0%	1 4.2% 0.7%	24 3.3%
70%	7 28.0% 8.0%	11 44.0% 5.0%	4 16.0% 2.2%	3 12.0% 3.1%	0 0.0% 0.0%	25 3.4%
80%	4 10.3% 4.6%	16 41.0% 7.3%	10 25.6% 5.6%	6 15.4% 6.1%	3 7.7% 2.1%	39 5.3%
90%	5 15.1% 5.7%	12 36.4% 5.5%	11 33.3% 6.2%	5 15.0% 5.1%	0 0.0% 0.0%	33 4.5%
100%	15 3.1% 17.1%	112 23.2% 50.9%	135 28.0% 75.4%	80 16.6% 81.6%	140 29.1% 96.6%	482 66.0%
Total and % all programs‖	88 12.1%	220 30.1%	179 24.5%	98 13.4%	145 19.8%	730 100.0%

*The authors analyzed data provided to them by the National Resident Matching Program to determine whether the students or the hospitals' specialty programs had greater success in the 1986 Match and to identify factors contributing to the success of each group. This table makes clear the observed trend that the more candidates a hospital listed for a specialty program position, the more likely was the position to be filled.

†The filling rate is the ratio between the number of positions filled and the number of positions offered by a program. In this table the programs are grouped by their filling rates, after these rates were rounded to the nearest 10%.

‡The top value in each group of three values = the number of programs; the middle value = the percentage that this number of programs is of all the programs with the designated filling rate; and the bottom value = the percentage that this number of programs is of all the programs with the designated number of candidates listed per position. For example, in the top left group, there were 20 programs; this was 62.5% of the 32 programs with the filling rate of 0% and was 22.7% of the 88 programs that listed 0-2.9 candidates per position.

§This is the total number of programs for the row (i.e., added horizontally), accompanied by the percentage that this number is of all the internal medicine programs studied.

‖This is the total number of programs for the column (i.e., added vertically), accompanied by the percentage that this number is of all the internal medicine programs studied.

Table 4

Impact of the Number of Choices of Internal Medicine Programs That Students Made on the Students' Success in Being Matched to a Program by the 1986 Match*

Level of Students' Choice Matched	Number of Choices Listed by Students					Total and
	1–3 No., %'s of Students†	4–6 No., %'s of Students†	7–9 No., %'s of Students†	10–12 No., %'s of Students†	13 or More No., %'s of Students†	No., %'s % All Students‡
1st	264 12.1% 52.9%	414 19.0% 66.8%	529 24.3% 57.3%	364 16.7% 50.8%	606 27.8% 47.4%	2,177 53.9%
2nd	27 3.8% 5.4%	87 12.4% 14.0%	188 26.7% 20.4%	167 23.7% 23.3%	235 33.4% 18.4%	704 17.4%
3rd	3 0.7% 0.6%	46 10.9% 7.4%	105 24.9% 11.4%	84 19.9% 11.7%	183 43.5% 14.3%	421 10.4%
4th	0 0.0% 0.0%	18 8.2% 2.9%	43 19.6% 4.7	44 20.1% 6.2%	114 52.1% 8.9%	219 5.4%
≥5th	0 0.0% 0.0%	7 3.0% 1.1%	44 19.0% 4.8%	50 21.6% 7.0%	131 56.5% 10.2%	232 5.8%
Unmatched	205 72.2% 41.1%	48 16.9% 7.7%	15 5.3% 1.6%	7 2.5% 1.0%	9 3.2% 0.7%	284 7.0%
Total and % all students§	499 12.4%	620 15.4%	924 22.9%	716 17.7%	1,278 31.7%	4,037 100.0%

*The authors analyzed data provided to them by the National Resident Matching Program to determine whether the students or the hospitals' specialty programs had greater success in the 1986 Match and to identify factors contributing to the success of each group. This table makes clear the observed trend that the more programs a student chose, the more likely was the student to be matched to a program.

†The top value in each group of three values = the number of students; the middle value = the percentage that this number of students is of all the students whose designated choice was matched; and the bottom value = the percentage that this number of students is of all the students who listed the designated number of program choices. For example, in the top left group, there were 264 students; this was 12.1% of the 2,177 students whose first choice was matched and was 52.9% of the 499 students who listed only one to three program choices.

‡This is the total number of students for the row (i.e., added horizontally), accompanied by the percentage that this number is of all the students who applied for internal medicine residency programs.

§This is the total number of students for the column (i.e., added vertically), accompanied by the percentage that this number is of all the students who applied for internal medicine residency programs.

fects the program's filing rate; here, we present more detailed data for one specialty, internal medicine. It is easy to see that those programs that did not fill any position — 20 out of total of 32 (62.5%) — listed fewer than three candidates per position. The percentage of the programs that filled all positions in each group significantly increased when the number of candidates listed per position increased (from 17.1% for the first group, with fewer than three candidates per position, to 96.6% for the fifth group, with 12 or more candidates per position).

Table 4 further illustrates the impact of the number of choices on the student's match; here we present more detailed data for the case of the U.S. senior students who applied for internal medicine programs only. The table shows that for those students who listed three of fewer choices, 205 out of 499 (41.1%) were unmatched. However, for those students who listed ten choices or more, fewer than 1% were unmatched. It is clear that listing too few choices does involve a high risk of being unmatched, both from the program's side and from the student's side.

Discussion

Our analysis strongly suggests that the actual results of the 1986 Match do not favor programs, a finding that is contrary to the perception of many investigators[5–7] and to the theoretical argument that the algorithm used for the matching process is biased in favor of programs.[4,8] Overall, in our study, the students performed better than did the programs. Furthermore, the analysis of each specialty program shows that in the majority of cases (20 out of 22 cases) the students' performance was better than or at least as good as that of the programs. The analysis was performed only for the outcome of the 1986 Match; however, we have no reason to believe that 1986 was different from previous or subsequent years. In order to test our conclusion, the same type of data and analysis could be used for the results of the Match in other years.

The NRMP has been praised as an equitable system of residency matching. The two goals of adopting a centralized matching mechanism were to bring order and fairness to a chaotic application process for internship and residency positions. There is general agreement that the first goal was achieved. With respect to the second goal — fairness — the evidence indicates that improvement is still called for. An ideal situation would be that the students' success would be equal to the programs' success and that no "sector" could, through superior analysis of publicly available information, through access to non-publicly-available information, or through any other mechanism, bring about superior performance. This issue is dealt with by the authors in more depth elsewhere.[10]

How else can we improve fairness in the system? Our findings in the present study do not supply the answer. We feel that a promising and practical way to reach this goal is to try to improve the assignment algorithm that is currently used. An example for such an effort to improve the algorithm by using the utility approach is described elsewhere by Mehrez and the present authors.[11]

Finally, we should notice that the matching algorithm, although being optimal in terms of stability,[2] does not guarantee to maximize the num-be expected to assign a student to a ber of matches,since no algorithm can position that is not mutually agreeable, as shown on the rank-order lists of both parties.[12] Every year we find that while some graduates fail to receive their internship positions, there are still vacancies in some hospitals. After the 1986 Match, for example, there were still 2,634 positions (14%) unfilled, and 5,221 graduates (24.4%) did not receive any position. Our analysis of the Match results suggests that for both the programs and students, listing more choices will increase their chances of finding a Match. Those who choose fewer programs could find themselves without a position in spite of their impression (which they might find to be wrong) of being highly ranked by a hospital where they had been interviewed.[7] The impact of the number of choices on the Match rate was also studied elsewhere by using computer simulation.[13] The results were similar to those presented in this paper.

Financial support for this study was provided by the National Resident Matching Program of the United States and the Natural Science and Engineering Research Council of Canada. The authors thank Dr. John S. Graettinger (who at the time of this study was executive vice president, NRMP) for his encouragement and valuable suggestions; Mr. E. Peranson and Mr. G. Spira (National Matching Service, Inc.), for providing NRMP data; and Dr. G. Norman for helpful suggestions. They also acknowledge the constructive comments made by three anonymous referees, which improved the quality of this paper.

References

1. Graettinger, J. S., and Peranson, E. The Matching Program. *N. Engl. J. Med.* **304**(1981):1163-1165.
2. Roth, A. E. The Evaluation of Labour Market for Medical Interns and Residents: A Case Study in Game Theory. *J. Political Economy* **92**(1984):991-1016.
3. *NRMP Directory.* Evanston, Illinois: National Resident Matching Program, 1980: 18-42.
4. Gale, D., and Shapley, S. L. College Admissions and the Stability of Marriage. *Am. Mathematical Monthly* **62**(1962): 9-15.
5. Martz, B. W. Residency Matching. *J. Med. Educ.* **58**(1983):498-499.
6. Polk, H. R. A Time for Reconsideration of Sequential Surgical Matching Programs. *Am. J. Surg.* **151**(1986):195-196.
7. Herman, T. A. Playing Fair in the NIRMP. *N. Engl. J. Med.* **302**(1979):1425.
8. Williams, K. J., Worth, V. P., and Wolff, J. A. An Analysis of the Resident Match. *N. Engl. J. Med.* **304**(1981):1165-1166.
9. Jordan, W. J. Success of Minority Applicants in the National Residency Matching Program. *J. Nat. Med. Assoc.* **78**(1986): 737-739.
10. Gafni, A., and Yuan, Y. Students' Performance vs Hospitals' Performance in the Labor Market for Medical Interns and Residents: Is the Market "Informationally Efficient"? *J. Health Econ.* **8**(1989): 353-360.
11. Mehres, A., Yuan, Y., and Gafni, A. Stable Solutions vs Multiplicative Utility Solutions for the Assignment Problem. *Operations Res. Lett.* **7**(1988):131-139.
12. Roth, A. E. On the Allocation of Residents to Rural Hospitals: A General Property of Two-Sided Matching Market. *Econometrica* **54**(1986):425-427.
13. Yuan, Y., and Kochen, M. Simulation of Matching under Incomplete Information. In *Proceedings of International ASME [Advancement of Modelling and Simulation Techniques in Enterprises] Conference*, pp 61-72. Williamsburg, Virginia: 1 (1986).

Variation in Lengths of Training among Residents Enrolled in Pediatrics and Internal Medicine Training Programs

ROBERT A. HOEKELMAN, M.D., and RUTH M. PARKER, M.D.

Abstract—Data were gathered in 1987 from 180 pediatrics and 302 internal medicine residency training programs about (1) the length of vacation provided their residents, (2) time allowed them for absence from work for reasons other than vacation, (3) how many of them during the previous three years had been required to extend their training to make up absences from work, and (4) whether a system was in place to cover their unexpected absences from work. All these variables affect the duration — but not necessarily the quality — of training that residents actually experience. The data show that these variables demonstrated wide ranges that depended on the postgraduate year of training, the discipline, the program size, and the program type (university, university-affiliated, freestanding, or military). The Accreditation Council for Graduate Medical Education, training program directors, and the American Board of Pediatrics and the American Board of Internal Medicine each need to consider these variables in making their respective decisions about accrediting training programs, verifying the clinical competence of trainees, and certifying program graduates. *Acad. Med.* 65(1990):257-266.

It is assumed that physicians in training within various medical disciplines have similar educational experiences in terms of numbers and types of clinical exposures. To this end, the Accreditation Council for Graduate Medical Education (ACGME) sets and applies minimal educational standards for training through residency review committees established for each discipline. In this way, the ACGME accredits each training program periodically, usually every five to seven years. The educational standards applied relate to the program's institutional affiliation, scope of training, curriculum, learning experiences, staff, facilities, and size, and the evaluation procedures used to measure resident and faculty performances. Accreditation is essential for training programs in order to attract trainees wishing to become certified in their medical disciplines. As one of the criteria for certification, the various specialty boards require completion of a specified number of years of training in accredited programs. Training program directors attest to the specialty board that the training of each resident has or has not been completed satisfactorily.

Despite all this effort to set standards, there are differences in the educational experiences provided to residents. These differences depend on (1) the sizes and locations of training programs, which dictate the numbers and types of patients seen, the frequencies of night and weekend call, and the quantity, variety, and quality of faculty available for teaching; and (2) program policies that relate to the durations of training residents actually experience — that is, policies that govern the length of vacation provided each year and the amount of time allowed for absence from work due to sickness and other reasons that the resident is not required to make up through extended training time.

The purpose of the study reported in this article was to investigate how much variation existed in the lengths of the vacations and allowable leaves of absence without required make-up time provided by pediatrics and internal medicine residency training programs in the United States.

Method

In January 1987, an eight-item, one-page questionnaire was sent to the directors of all 657 accredited pediatrics and internal medicine residency training programs located in the United States and listed in the 1986-1987 Directory of Graduate Medical Education Programs.[1] (The questionnaire was not sent to the directors of the 15 listed programs located in Puerto Rico.) The questionnaire queried the directors on (1) whether the program was a university program, a university-affiliated program, a freestanding program, or a program located in a military hospital; (2) the number of residents currently enrolled in each of the four post-graduate years (PGY-1, PGY-2, PGY-3, and PGY-4); (3) the number of weeks of vacation allowed for each level of trainee; (4) the program's policy concerning making up absences from work (excluding vacation); (5) whether records of days absent from work were kept and, if so, who kept them; (6) the numbers of residents who had completed three years of training by June 1984, June 1985, and June 1986 who had had absences from work (excluding vacations) that exceeded three months; (7) the numbers of residents who had completed three years of training by June 1984, June 1985, and June 1986 who had been required to make up time; and (8) whether a resident was regularly scheduled to provide coverage for unexpected house-officer absences from work (a sick-call resident).

Of the 225 pediatrics residency program directors queried, 180 (80%) responded to the initial mailing; 302 of

---

Dr. Hoekelman is professor and chairman, Department of Pediatrics, University of Rochester School of Medicine and Dentistry, Rochester, New York. Dr. Parker was a Robert Wood Johnson Clinical Scholar, University of Pennsylvania School of Medicine, when this study was conducted and currently is assistant professor, Department of Internal Medicine, Emory University School of Medicine, Atlanta, Georgia.

Correspondence should be addressed to Dr. Hoekelman, 601 Elmwood Avenue, PO Box 777, Rochester, NY 14642. Reprints will not be available.

Table 1

Numbers and Percentages of 180 Pediatrics and 302 Internal Medicine Residency Programs, by Type and Size, 1987*

	Program Size						Total, by Type	
	Small		Medium		Large			
Program Type	No.	%	No.	%	No.	%	No.	%
Pediatrics								
University	15	16.3	26	28.3	51	55.4	92	51.1
University-affiliated	8	27.6	16	55.2	5	17.2	29	16.1
Freestanding	26	55.3	18	38.3	3	6.4	47	26.1
Military	8	66.7	4	33.3	0	0.0	12	6.7
Total, by size	57	31.7	64	35.6	59	32.8	180	100.0
Internal medicine								
University	4	5.3	15	19.7	57	75.0	76	25.2
University-affiliated	22	33.3	28	42.4	16	24.2	66	21.8
Freestanding	69	46.9	52	35.4	26	17.7	147	48.7
Military	6	46.2	7	53.8	0	0.0	13	4.3
Total, by size	101	33.4	102	33.8	99	32.8	302	100.0

*In January 1987, the authors sent a questionnaire to the directors of all 657 U.S. pediatrics and internal medicine residency programs in order to determine how much variation existed in the lengths of the residents' vacations and allowable leaves of absence without required make-up time. The data above describe the programs that responded.

the 432 internal medicine residency program directors did so (70%). We did not use a second mailing to increase this response rate because it was considered sufficiently high to satisfy the purposes of the study and the percentages of responses from directors of the four program types — university, university-affiliated, freestanding, and military — were essentially the same as the percentages of distribution of those four program types within the total sample for each of the two disciplines studied (see Table 1).

The data were analyzed according to the variables of discipline, program type, and program size, by use of frequency distributions and cross tabulations.

The authors defined the four program types, for each of the two disciplines, as (1) a *university program* — a residency training program conducted primarily by a university department of pediatrics or internal medicine and located in a general hospital, a children's hospital, or a Veterans Administration hospital; (2) a *university-affiliated program* — a residency training program located in a community hospital, a children's hospital, or a Veterans Administration hospital *not* conducted primarily by a university department of pediatrics or internal medicine, but whose residents are recruited and enrolled as students in a university's graduate medical education program; (3) a *freestanding program* — a residency training program located in a community hospital, a children's hospital, or a Veterans Administration hospital that has no affiliation with a university with respect to conducting graduate medical education programs for residents; and (4) a *military program* — a residency training program located in a U.S. Army, Air Force, or Navy hospital that has no affiliation with a university with respect to conducting graduate medical education programs for residents. These groupings differ from those used by the Accreditation Council on Graduate Medical Education (ACGME) when classifying participants in residency training programs by their relationship to a medical school.

There were 35 children's hospitals in the sample — 31 were university programs, one was university-affiliated, and three were freestanding. Thirty (86%) of these hospitals responded to the questionnaire. There were 13 Veterans Administration hospitals in the sample — three were university programs, five were university-affiliated, and five were freestanding. Ten (77%) of these hospitals responded to the questionnaire. Data from the children's and Veterans Administration hospital respondents were included with those of each of the categories of training programs they represented, rather than in separate children's and Veterans Administration hospital categories.

Program size for each of the two disciplines was determined by categorizing the programs that responded into three groups (small, medium-sized, and large) according to the number of residents enrolled. Thus, for the 180 pediatrics programs the numbers of residents enrolled ranged from five to 110. Fifty-seven of the programs had 15 or fewer residents enrolled and were designated small programs, 64 had between 16 and 34 residents enrolled and were designated medium-size programs, and 59 had 35 or more residents enrolled and were designated large programs. (Notice that the numbers of programs in the groups are approximately equal.) For the 302 internal medicine programs, the numbers of residents enrolled ranged from four to 159. A total of 101 of the programs had 26 or fewer residents enrolled (small programs), 102 had between 27 and 46 residents (medium-sized programs), and 99 had 47 or more residents (large programs). Here again, the numbers of programs in the groups are about the same. Response rates for the three program sizes within the entire sample could not be determined because the numbers of residents enrolled in the programs that did not respond were not known. Table 1 shows the numbers and percentages of responding programs by program type and size.

Results

The data obtained from the responding programs are presented here in terms of the number of residents enrolled, the length of vacations provided to them, the other absences from work allowed to them and experienced by them, and the use of a sick-call resident system to cover them when they were unexpectedly absent from work.

Number of Residents in Training

The residents enrolled in all the responding pediatrics programs during the 1986–87 academic year were fairly evenly distributed among the first three training years (1,734 in PGY-1, 1,642 in PGY-2, and 1,568 in PGY-3), showing no strong tendency towards the "pyramid" effect—that is, significant losses between PGY-1 and PGY-3. There were 218 PGY-4 residents enrolled; 25% of the 180 pediatrics programs had one PGY-4 resident and 26% had two or more.

On the other hand, the responding internal medicine programs enrolled 23% fewer PGY-2 and PGY-3 residents than PGY-1 residents (3,764 from PGY-2 and 3,745 from PGY-3 versus 4,910 from PGY-1). This reflects the enrollment of "transitional-year" residents who enter other, non-internal-medicine, training programs following their first postgraduate year. There were 436 PGY-4 residents enrolled; 23% of the 302 internal medicine programs had one PGY-4 resident, and 34% had two or more.

Table 2 shows (1) the number of residents enrolled in the responding pediatrics and internal medicine programs at all levels of training, by each program type and size; (2) the per-

Table 2

Numbers and Percentages of Residents Enrolled in 180 Pediatrics and 302 Internal Medicine Programs, by Program Type and Size, 1986–87*

Program Type and Size	No. Programs	No. Residents	% All Residents, by Program Type	% All Residents in All Programs
Pediatrics				
University				
Small	15	177	5.1	3.4
Medium	26	661	19.1	12.8
Large	51	2,616	75.7	50.7
Total	92	3,454	99.9†	66.9
University-affiliated				
Small	8	101	14.0	1.9
Medium	16	387	53.8	7.5
Large	5	231	32.1	4.5
Total	29	719	99.9†	13.9
Freestanding				
Small	26	300	36.3	5.8
Medium	18	380	46.0	7.4
Large	3	146	17.7	2.8
Total	47	826	100.0	16.0
Military				
Small	8	96	58.9	1.9
Medium	4	67	41.1	1.3
Large	0	—	—	—
Total	12	163	100.0	3.2
All programs	180	5,162	—	100.0
Internal medicine				
University				
Small	4	86	1.6	0.7
Medium	15	536	10.2	4.2
Large	57	4,630	88.2	36.0
Total	76	5,252	100.0	40.9
University-affiliated				
Small	22	410	15.7	3.2
Medium	28	1,092	41.8	8.5
Large	16	1,109	42.5	8.6
Total	66	2,611	100.0	20.3
Freestanding				
Small	69	1,261	27.3	9.8
Medium	52	1,818	39.3	14.1
Large	26	1,543	33.4	12.0
Total	147	4,622	100.0	35.9
Military				
Small	6	115	31.1	0.9
Medium	7	255	68.9	2.0
Large	0	—	—	—
Total	13	370	100.0	2.9
All programs	302	12,855	—	100.0

*In January 1987, the authors sent a questionnaire to the directors of all 657 U.S. pediatrics and internal medicine residency programs in order to determine how much variation existed in the lengths of the residents' vacations and allowable leaves of absence without required make-up time. The above data describe the programs that responded.

†This total does not equal 100% because the component percentages were rounded.

Table 3

Numbers and Percentages of 180 Pediatrics and 302 Internal Medicine Residency Programs with Specified Vacation Lengths, by Level of Training and Program Type, 1987*

Program Type for Each Postgraduate Year	No. Programs	Percentage of Programs with Specified Length of Vacation		
		Two Weeks	Three Weeks	Four Weeks
Pediatrics				
PGY-1				
University	92	34.7	32.6	32.6
University-affiliated	29	34.5	34.5	31.0
Freestanding	47	27.8	34.0	38.2†
Military	12	100.0	0.0	0.0
All programs	180	36.7	31.7	31.7
PGY-2				
University	92	7.6	48.9	43.5
University-affiliated	29	3.4	51.7	44.8
Freestanding	47	6.4	48.9	44.6†
Military	12	50.0	16.7	33.3
All programs	180	9.4	47.8	42.8
PGY-3				
University	92	4.3	47.8	47.8
University-affiliated	29	3.4	41.4	55.2
Freestanding	47	6.4	42.6	51.0†
Military	12	33.3	16.7	50.0
All programs	180	6.7	43.9	49.4
Internal medicine				
PGY-1				
University	76	44.7	32.9	22.4
University-affiliated	66	39.4	37.9	22.7
Freestanding	147	40.8	38.1	21.1
Military	13	100.0	0.0	0.0
All programs	302	44.0	35.1	20.9
PGY-2				
University	76	9.2	56.6	34.2
University-affiliated	66	7.6	56.1	36.3
Freestanding	147	12.9	58.5	28.6
Military	13	38.5	38.5	23.0
All programs	302	11.9	56.6	31.5
PGY-3				
University	76	5.3	59.2	35.5
University-affiliated	66	3.0	56.1	40.9
Freestanding	147	10.2	54.4	35.4
Military	13	30.8	23.1	46.1
All programs	302	8.3	54.6	37.1

*In January 1987, the authors sent a questionnaire to the directors of all 657 U.S. pediatrics and internal medicine residency programs in order to determine how much variation existed in the lengths of the residents' vacations and allowable leaves of absence without required make-up time. The above data present the vacation lengths of the programs that responded.

†One medium-sized, freestanding program gave five weeks of vacation in each of the four years of training.

centage of all residents enrolled, by each program type and size; and (3) the percentage of all residents these numbers represent for all the pediatrics and internal medicine residency training programs that responded. Thus, 66.9% of all these pediatrics residents were enrolled in the university programs, while only 40.9% of the internal medicine residents received their training in such programs. Freestanding programs accounted for 35.9% of the internal medicine trainees and 16% of the pediatrics trainees. The university-affiliated programs were lower on the list, accounting for 20.3% of the internal medicine and 13.9% of the pediatrics trainees. The military hospitals trained relatively few residents—approximately 3% for each discipline.

Length of Vacation

The length of vacation that residents received during each of the three years of training varied by their levels of training, their disciplines, and the types and sizes of the programs in which they were enrolled. Tables 3

Table 4

Numbers and Percentages of 180 Pediatrics and 302 Internal Medicine Residency Programs with Specified Vacation Lengths, by Level of Training and Program Size, 1987*

Program Size for Each Postgraduate Year	No. Programs	Percentage of Programs with Specified Length of Vacation		
		Two Weeks	Three Weeks	Four Weeks
Pediatrics				
PGY-1				
Small	57	54.4	21.1	24.5
Medium	64	32.8	31.3	35.9†
Large	59	23.7	42.4	33.9
All programs	180	36.7	31.7	31.7
PGY-2				
Small	57	19.3	49.1	31.6
Medium	64	7.8	43.8	48.4†
Large	59	1.7	50.8	47.5
All programs	180	9.4	47.8	42.8
PGY-3				
Small	57	12.3	50.9	36.8
Medium	64	7.8	34.4	57.8†
Large	59	0	47.5	52.5
All programs	180	6.7	43.9	49.4
Internal medicine				
PGY-1				
Small	101	52.5	41.6	5.9
Medium	102	45.1	33.3	21.6
Large	99	34.3	31.3	34.3
All programs	302	44.0	35.1	20.9
PGY-2				
Small	101	16.8	66.3	16.8
Medium	102	10.8	58.8	30.4
Large	99	8.1	45.5	46.5
All programs	302	11.9	56.6	31.5
PGY-3				
Small	101	9.9	68.3	21.8
Medium	102	8.8	52.9	38.2
.Large	99	7.1	42.4	50.5
All programs	302	8.3	56.6	37.1

*In January 1987, the authors sent a questionnaire to the directors of all 657 U.S. pediatrics and internal medicine residency programs in order to determine how much variation existed in the lengths of the residents' vacations and allowable leaves of absence without required make-up time. The above data present the vacation lengths of the programs that responded.

†One medium-sized freestanding program gave five weeks of vacation in each of the four years of training.

and 4 demonstrate these variations. The data presented in the tables show that in general (1) vacation time increased each year; (2) the pediatrics residents got more vacation time at all levels than did the internal medicine residents; (3) the university-affiliated and freestanding pediatrics programs provided more vacation time to their residents than did the university and military pediatrics programs; (4) the university-affiliated internal medicine programs provided more vacation time to their residents than did the other types of internal medicine programs; and (5) the larger the program size, the longer the vacation time provided to residents in both disciplines, except that among the pediatrics programs, the medium-sized ones provided longer vacations during each year of training than did the large ones.

During the 1986–87 academic year, less than half of all the programs in both disciplines provided four weeks of vacation to their residents at any of the levels of training. Indeed, only 37.1% of the internal medicine programs allowed four weeks of vacation for their residents in PGY-3; those in PGY-2 and PGY-1 fared less well (31.5% and 20.9%, respectively). As noted above and shown in Tables 3 and 4, the pediatrics programs were more generous in this respect. At the PGY-4 level of training (not shown in the tables), 57% of the pediatrics and 44% of the internal medicine programs allowed four weeks of vacation. Eight of the pediatrics and six of the internal medicine programs that provided two to three weeks of vacation for their PGY-2 and PGY-3 residents reported that they allowed them one to two weeks each year to attend medical meetings or for independent study.

Absences from Work

The number of weeks, in addition to vacation time, that the pediatrics and internal medicine training programs allowed their residents to be absent from work because of illness, maternity leave, paternity leave, or family crises (illness, death, or discord) varied considerably by program type and size both within each discipline and between the two disciplines, as shown in Table 5. Only 6.7% of the pediatrics programs and 7.6% of the internal medicine programs allowed no such absences, requiring residents to make up *all* time missed from work. Some programs (17.2% of the pediatrics and 20.5% of the internal medicine programs) had no policy governing these absences. These respondents indicated that absences from work beyond vacation time were allowed without residents' having to make up the time, but that the amount of time allowed depended on (1) the individual circumstances surrounding leaves for sickness, maternity, paternity, or family crisis; (2) the number of consecutive weeks absent from work; (3) the training year during which the absence occurred—that is, absences during the PGY-1 year were more likely to require make-up than those occurring during PGY-2 or PGY-3; (4) the specific service to which the resident was assigned during the absence—that is, absences during intensive care unit assignments were more likely to require make-up than those occurring during other assignments; and (5) the resident's overall performance, with high-performing residents being less likely to be required to make up time away from work than poorly-performing residents.

Of the 180 pediatrics residency programs responding, 127 (70.6%) stated that they kept records of the times residents were absent from work. The university programs were least likely (58.7%) and the military programs most likely (91.7%) to do so. The university-affiliated and the freestanding programs kept such records at rates of 79.3% and 82.9%, respectively. The small and medium-sized programs were most apt to keep such records (77.2% and 78.1%) and the large programs least apt to (55.9%).

Similar distributions for keeping records of residents' absences were reported by the 302 internal medicine programs—71.2% of all the programs, 57.3% of the university programs, 80.3% of the university-affiliated programs, and 77.1% of the freestanding programs. However, only 61.5% of the military programs did so. The small internal medicine programs kept records of residents' absences most frequently (84.8%), while the medium-sized and large programs did so less frequently (65.7% and 64.6%, respectively).

The responsibility for keeping absence-from-work records in the 482 pediatrics and internal medicine programs fell to the program chairman's office (77.2% of all programs), to the chief resident (14.7% of all programs), or to the individual resident (8.1% of all programs). There was little variation in assignments of this responsibility among the disciplines or by program type or size within them. Thirteen of the respondents volunteered that while they kept records of residents' absences, they did so only for absences that exceeded one week.

Of the approximately 4,700 pediatrics residents (enrolled in the 180 responding programs) who were scheduled to complete three years of training on June 30 in 1984, 1985, or 1986, 94 (2%) were required to make up time missed from work. For the same period, 136 (1.7%) of the approximately 8,200 internal medicine residents enrolled in the 302 responding programs were required to do so. Neither the lengths of make-up time required in these instances nor whether this time was taken from unused vacation periods or added onto the end of the three years of training was determined through the questionnaire. There was no tendency in relation to program type or size, or over the three years surveyed, to require make-up time more frequently or less frequently.

There were 57 pediatrics and 63 internal medicine programs that provided 12 weeks of vacation over the three years of training. Of these, 22 pediatrics and 28 internal medicine programs had policies that allowed up to 15 extra weeks of absence from work for other reasons during the three years of training, without requiring make-up. Thus, any illness, maternity, paternity, or family crisis that led to any days away from work in these programs caused their residents to exceed three months of absence from work during the three-year training period. This is the extreme circumstance, but one that must have affected large numbers of residents. Yet only 31 pediatrics and 12 internal medicine residents enrolled in these 50 programs who completed their three years of training at the end of June in 1984, 1985, or 1986 were required to make up time. There surely were many more who were not required to do so, as there must have been among those enrolled in programs that provided less than 12 weeks of vacation, but who were absent from work more than three months overall. They must have numbered many more than the 94 pediatrics and 136 internal medicine residents (of the over 18,000 enrolled annually in all of the responding programs) who were required to make up time either during or following the three years of training that ended on June 30, 1984, 1985, or 1986. Indeed, 28 of the pediatrics and 48 of the internal medicine program directors spontaneously responded that they would allow residents from three weeks to three months of absence from work beyond vacation time during the three years of training without requiring that this time be made up.

The specific issue of the amount of time that each program allowed for

Table 5

Numbers and Percentages of 180 Pediatrics and 302 Internal Medicine Residency Programs with Specified Weeks of Absence Allowed Residents (beyond Vacation Time) during Three Years of Training without Requiring Make-Up Time, by Program Type and Size, 1987*

Program Type and Size	No. Programs	Percentage of Programs with Specified Weeks of Absence Allowed					
		None	1–2	3–4	5–6	7–15	No Policy
Pediatrics							
University							
Small	15	20.0	0.0	6.7	13.3	26.7	33.3
Medium	26	7.7	7.7	15.4	15.4	26.9	26.9
Large	51	7.8	5.9	17.6	13.7	39.2	15.7
All programs	92	9.8	5.4	15.2	14.1	33.7	21.7
University-affiliated							
Small	8	12.5	0.0	12.5	12.5	50.0	12.5
Medium	16	0.0	6.3	31.3	0.0	50.0	12.5
Large	5	0.0	20.0	20.0	0.0	60.0	0.0
All programs	29	3.4	6.9	24.1	3.4	51.7	10.3
Freestanding							
Small	26	0.0	11.5	26.9	15.4	34.6	11.5
Medium	18	5.5	5.5	27.8	16.7	27.8	16.7
Large	3	33.3	0.0	66.7	0.0	0.0	0.0
All programs	47	4.3	8.5	29.8	14.9	29.8	12.8
Military							
Small	8	0.0	25.0	37.5	12.5	12.5	12.5
Medium	4	0.0	0.0	0.0	25.0	50.0	25.0
Large	0	—	—	—	—	—	—
All programs	12	0.0	16.7	25.0	16.7	25.0	16.7
All pediatrics programs	180	6.7	7.2	21.1	12.8	35.0	17.2
Internal medicine							
University							
Small	4	0.0	0.0	25.0	25.0	0.0	50.0
Medium	15	13.3	6.7	13.3	13.3	20.0	33.3
Large	57	12.3	3.5	24.6	8.8	35.1	15.8
All programs	76	11.8	3.9	22.4	10.5	30.3	21.1
University-affiliated							
Small	22	13.6	4.5	18.2	18.2	22.7	22.7
Medium	28	7.1	10.7	21.4	10.7	21.4	28.6
Large	16	0.0	12.5	12.5	12.5	25.0	37.5
All programs	66	7.6	9.1	18.2	13.6	22.7	28.8
Freestanding							
Small	69	13.0	8.7	23.2	8.7	26.1	20.3
Medium	52	0.0	9.6	26.9	17.3	30.8	15.4
Large	26	0.0	0.0	34.6	7.7	42.3	15.4
All programs	147	6.1	7.5	26.5	11.6	30.6	17.7
Military							
Small	6	0.0	33.3	50.0	0.0	16.7	0.0
Medium	7	0.0	14.3	14.3	14.3	42.8	14.3
Large	0	—	—	—	—	—	—
All programs	13	0.0	23.1	30.8	7.7	30.8	7.7
All internal medicine programs	302	7.6	7.6	23.8	11.6	28.8	20.5

*In January 1987, the authors sent a questionnaire to the directors of all 657 U.S. pediatrics and internal medicine residency programs in order to determine how much variation existed in the lengths of the residents' vacations and allowable leaves of absence without required make-up time. The above data present the lengths of the allowable weeks of absence, if any, of the programs that responded.

maternity leave was not addressed in the questionnaire. However, nine program directors commented that they had had one resident over the prior three years who had had to make up time taken away from work because of pregnancy. On the other hand, three program directors commented that their programs allowed between 30 and 60 days (in addition to vacation time) for maternity leave that did not have to be made up. One of these programs also allowed seven days for paternity leave that did not require make-up.

Use of a Sick-Call Resident System

A sick-call resident system provides coverage for residents who are unexpectedly absent from work for any reason.[2] Residents are assigned to this function within the program's schedule, usually while they are on a subspecialty elective. The system ordinarily provides coverage for night and weekend duties, although in some programs week-day assignments are also covered by the sick-call resident. The system is designed to reduce stress among residents who must assume the added work and on-call hours created by a peer's unexpected absence and to reduce stress among the residents who are absent from work—stress caused by the reason for the absence itself and the guilt of adding to their colleagues' already heavy work load.

A sick-call resident system was used in 37 (20.6%) of the 180 responding pediatrics programs. Thirty of the 92 university programs (32.6%), five of the 29 university-affiliated programs (17.2%), two of the 47 freestanding programs (4.3%), and none of the 12 military hospital programs had a sick-call resident system. Large programs were more likely to have such a system (30.5%) than were medium-sized programs (21.9%) or small programs (8.8%).

Of the 302 responding internal medicine programs, 98 (32.5%) had a sick-call resident system. Thirty-two of the 76 university programs (42.7%), 24 of the university-affiliated programs (36.4%), 42 of the 147 freestanding programs (29.2%), and none of the 13 military hospital programs had a sick call resident system. Again, large programs were more likely to have such a system (45.5%) than were medium-sized (34.3%) or small programs (17.8%).

Those pediatrics and internal medicine programs without a sick-call resident system either used the chief resident of hired part-time physicians to cover for residents who were sick or absent from work for other reasons, or had no backup system at all. In the latter instance, the extra work load was assumed by residents assigned to the same service as the absent resident.

Discussion

The data presented in the previous section are discussed in relation to (1) requirements, promulgated by the pediatrics and internal medicine residency review committees of the Accreditation Council for Graduate Medical Education, for accreditation of training programs; (2) training requirements that applicants to the American Board of Pediatrics and the American Board of Internal Medicine must meet for certification; and (3) recent attempts to reduce the stress experienced by pediatrics and internal medicine residents during training.

The special requirements for residency training in pediatrics[3] and in internal medicine[4] generated by their respective residency review committees make no mention of the length of vacation that should be provided residents or of restrictions on leaves of absence for any reason during training. The residency review committee for internal medicine, however, does make specific recommendations regarding work schedules, which were implemented on October 1, 1989. These include (1) being on call every third or fourth night during inpatient rotations and on call no more than every third night, regardless of the assignment; (2) working no more than 80 hours per week, averaged over four weeks; (3) being free of hospital duties at least one day out of seven; and (4) drawing assignments to the emergency department that do not exceed 12 continuous hours and are followed by at least 8 hours off duty. Recommendations 2, 3, and 4 are identical to requirements implemented by the New York State Department of Health on July 1, 1989. New York State's limitation that residents can work no more than 24 consecutive hours in direct patient care activities when assigned to inpatient services is not included among the recommendations of the internal medicine residency review committee. The pediatrics residency review committee currently makes no recommendations to limit residents' work schedules, but one of its revised requirements (effective January 1, 1990) is that "resident on-call rotations should occur with an average frequency of every third or fourth night in any year of the program and the schedule should be designed to provide a monthly average of at least one day out of seven without assigned duties in the program."[3] In a 1985 memorandum, the residency review committee stated, that "absenteeism (illness, maternity leave, etc.) which exceeds three months in three years, excluding approved vacation, could adversely affect the attainment of educational objectives, except in very unusual circumstances."[5]

The American Board of Pediatrics, on the other hand, requires at least 33 months of training for residents. In its booklet of information for applicants it states that "absences in excess of three months, whether for vacation, sick leave, maternity leave, etc., must be made up."[6] Further, the board requires residency program directors, in completing its Verification of Clinical Competence Form for each applicant, "describe any periods of absence from the training program that have not been made up by the applicant. This would include any absences, maternity/paternity leaves or sick leave taken in addition to the vacation and sick leave normally allowed each resident. The Board recommends that absences from a program beyond three months during the three years of training be made up by additional periods of training."[7] If the program director believes that an absence of more than three months is justified and that the candidate is qualified to take the board examination, a letter of explanation can be sent by the director to the board's credentials committee for review.[8]

The American Board of Internal Medicine makes no statement in its Policies and Procedures booklet that limits the amount of absence from work that residents may be permitted during their three yeas of training.[9] It explains in a November 19, 1984, memorandum to program directors that this reflects a flexible policy of the board concerning acceptable reductions in its training requirements and leaves to the discretion of program directors the length of absences from work that is permitted residents without requiring make-up. Justification of absences greater than one month per year, exclusive of vacation time and acceptable to the board, is expected from the program director.

The American Board of Pediatrics' posture regarding the length of absences from work permitted residents without requiring make-up presents difficulties for both residents and program directors. The data presented in Tables 3 and 4 show that 31.7% of the responding pediatrics programs allowed 12 weeks of vacation during the three years of training; by board criteria, this would allow no time for absences due to illness or maternity, paternity, or family crises without requiring make-up or getting program-director justification for such absences. The weeks of absence allowed without having to be made up by most of the rest of the programs, shown in Table 5, indicate that large numbers of pediatrics residents were not meeting American Board of Pediatrics length-of-training requirements during the period surveyed in this study. If program directors reported all residents' absences from work to the board on its Verification of Clinical Competence Form, most of these residents would be required to extend the length of their training beyond three calendar years. Program costs for resident stipends, fringe benefits, and malpractice insurance would rise appreciably, and residents would be delayed in starting practice or subspecialty fellowship training. It would also interfere with the timing for spouses' training or job placements in other communities. This could explain why only 2% of the pediatrics residents were required to make up time at the end of their training over the three years surveyed.

The stress caused by worrying about having to make up time, in addition to the already high levels of stress residents experience during training, should be of great concern to program directors. Methods for reducing these high levels of stress in pediatrics and internal medicine programs have been recommended by many individuals and groups involved in directing residency training. The most recently published recommendations emphasize the importance of flexibility in scheduling residents' work hours and in allowing adequate vacation time and absence from work when needed for other reasons.[10,11] It is disappointing that so many programs allow their residents only two or three weeks of vacation each year and require invasion of this time to make up for other absences, and that so few programs have implemented a sick-call resident system.

As noted in the introduction to this report, there are many variables that influence the quality of residency training; the amount of time spent at work over the three years of training may be the least of these. The quality of training programs is the focus of residency review committees, and those for pediatrics and internal medicine have focused on the training program's structure (institution and faculty), process (curriculum and clinical experiences), and outcome (success of its trainees on board-certifying examinations) in judging a program's quality. Other outcomes of training, such as the quality of medical care former trainees deliver, are difficult to measure and are influenced by intervening variables not related to the quality of training received. The residency review committees have chosen not to focus on duration of training, except for the number of years spent at it, perhaps recognizing that even this is arbitrary and that the learning gained during one week of training in one program is equal to that gained during two or three weeks in another. Rather, they have depended heavily on program directors to determine when their trainees are ready to "come out."

Nevertheless, the problem of wide variations in the length of training experienced by pediatrics and internal medicine residents and its effect upon the quality of their education and their lives needs to be resolved.

There are no easy solutions. Standardizing vacations to allow four weeks during each year of training would reduce some of the variation in length of training and the levels of stress among residents, but this would be difficult for small programs to implement without increasing the work load of residents not on vacation or without reducing the number of patients covered by the residents. Standardizing the length of absence from work allowed for other reasons (for example, to ten, 20, or 30 days during each year of training or, over the three years, to 30, 60, or 90 days) would present the same advantages and disadvantages. In addition, some program directors, residency review committee members, and specialty board members would be concerned that the overall duration of training for substantial numbers of residents would be shortened at a time when the number of residency review committee–required clinical experiences are increasing, and that some residents would take advantage of the system to maximize their time away from work.

It would seem most appropriate for the boards of pediatrics and of internal medicine to adopt flexible policies that transfer to program directors the responsibility to determine within reasonably broad limits how much time each resident should spend at work during each year of training. However, program directors would have to do this diligently and with the greatest of care and wisdom.

The authors thank the 482 pediatrics and internal medicine residency training program directors who provided the information upon which this report is based.

References

1. *1986–1987 Directory of Graduate Medical Education Programs.* Chicago, Illinois: American Medical Association, 1986.

2. Parker, R. M., Hoekelman, R. A., and Napodano, R. J. Illnesses and Other Causes of Unexpected Absences From Work During Residency Training *J. Med. Educ.* **66**(1987):959–968.
3. *Revised Requirements for Residency Training in Pediatrics.* Chicago, Illinois: Accreditation Council for Graduate Medical Education. [Fourteen-page memorandum to directors of pediatrics residency programs] February 18, 1989.
4. *Special Requirements for Residency Training Programs in Internal Medicine.* Chicago, Illinois: Accreditation Council for Graduate Medical Education. [Fifteen page memorandum to directors of internal medicine residency programs] September 27, 1988.
5. *Revised Special Requirements.* Chicago, Illinois: Accreditation Council for Graduate Medical Education. [Ten-page memorandum to directors of pediatrics residency programs] February 19, 1985, p. 9.
6. *Booklet of Information.* Chapel Hill, North Carolina: The American Board of Pediatrics. 1988, p. 16.
7. *Verification of Clinical Competence Forms.* Chapel Hill, North Carolina: The American Board of Pediatrics. [Four-page memorandum to pediatrics program directors] March 21, 1988.
8. Oliver, T. K., Jr., Senior Vice-President, American Board of Pediatrics. Personal communication. August 30, 1989.
9. *Policies and Procedures.* Philadelphia, Pennsylvania: American Board of Internal Medicine, 1988.
10. Hoekelman, R. A. Stress Experienced during Pediatric Residency Training: Its Causes, Consequences, Recognition, and Solutions. *Am. J. Dis. Child.* **143**(1989): 177–180.
11. Resident Services Committee, Association of Program Directors in Internal Medicine. Stress and Impairment during Residency Training: Strategies for Reduction, Identification and Management. *Ann. Intern. Med.* **109**(1988):154–161.

Surveying Graduates of Combined Internal Medicine–Pediatrics Residency Programs

FRANK M. BIRO, M.D., MATTHEW W. GILLMAN, M.D., RUTH M. PARKER, M.D., PHILIP R. KHOURY, M.S., and DAVID M. SIEGEL, M.D.

Abstract—Graduates of all U.S. combined internal medicine-pediatrics residency programs were surveyed in 1987 regarding a variety of demographic information about their residencies and current practices, the residency curricula they had followed for both specialties, and recommendations for modifications in training. The 71 responding graduates (from a total of 112) reported patient care as their major involvement (mean of 42.9 hours per week), with a majority (83%) seeing patients in both pediatric and adult age groups. Most were involved in primary care only (64%). The graduates reported that during both pediatrics and internal medicine training, they had had too many inpatient and intensive care rotations and too few elective and ambulatory rotations. The most important subspecialty rotations in internal medicine were considered to be cardiology, dermatology, and pulmonary medicine; and in pediatrics, infectious disease, cardiology, and adolescent medicine. The graduates recommended more outpatient subspecialty rotations, ambulatory rotations in medicine and pediatrics, and a combined medicine-pediatrics continuity clinic. *Acad. Med.* 65(1990):266–271.

In the past decade, combined internal medicine and pediatrics residency training programs (med-ped) have become increasingly common. Originally evolved in the 1960s from a two-year rotating internship,[1] med-ped programs now are often viewed as an alternative to family-practice training. The programs are generally four years in length and divided equally between the two specialties. By virtue of a 20-year-old agreement between the American Board of Pediatrics and the American Board of Internal Medicine, graduates of such residencies are eligible for both board examinations.

In 1980 there were four med-ped programs identified in the National Resident Matching Program (NRMP) Directory, with nine available postgraduate year-one positions. Only 20 physicians had completed the combined residency by that time. By 1988 there were 76 active programs listed by the NRMP.[2] In 1980, the NRMP reported only 20 applications for med-ped programs. By 1987, there were 1,110 applications from American senior medical students, which is comparable to the 1,755 applications in primary care internal medicine in the same year. This is in contrast to 28,941 applications from American students for all categories of internal medicine programs in 1980, which dropped to 27,239 by 1987.[3] Little is known about the careers med-ped graduates actually choose: whether the graduates are academia- or community-based, whether they are primary or subspecialty providers, and how they are involved with various age groups.[4,5] Questions also remain regarding optimal curricula for trainees, how graduates are accepted by their peers in categorical residencies, and how they perceive their competencies in representative clinical situations. In order to address these questions concerning internal medicine–pediatrics training pro-

---

Portions of these data were presented at the annual meeting of the Society for Pediatric Research, Washington, D.C., on May 5, 1988, and at the annual meeting of the Society of General Internal Medicine, Washington, D.C., on April 29, 1988.

Dr. Biro is assistant professor, and Mr. Khoury is a biostatistician, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio; Dr. Gillman is staff physician, South Boston Community Health Center, Boston, Massachusetts; Dr. Parker is assistant professor of medicine, Department of Internal Medicine, Emory University School of Medicine, Atlanta, Georgia, and was a fellow in the Robert Wood Johnson Clinical Scholars Program at the University of Pennsylvania School of Medicine in Philadelphia at the time of this study; and Dr. Siegel is assistant professor, Department of Pediatrics, University of Rochester School of Medicine and Dentistry, Rochester, New York.

Correspondence and requests for reprints should be addressed to Dr. Biro, Division of Adolescent Medicine, Children's Hospital Medical Center, Elland and Bethesda Avenues, Cincinnati, OH 45229.

grams, the investigators performed the first nationwide survey of graduates of such combined residencies; this article is a report of their investigation and what they found.

Method

Through telephones and written inquiries to the directors of all the med-ped programs listed by the NRMP in 1987, the investigators attempted to obtain the addresses of all the graduates of these programs who had completed combined residencies before July 1986. Questionnaires were mailed directly to the 88 graduates of ten of the 16 programs. The directors of the remaining six programs chose to have their offices mail the surveys to their graduates (an additional 24), citing their policies regarding privacy of information to explain why they did not disclose the names and addresses of these graduates. The questionnaire covered several areas: demographics, residency format, suggested modifications of training, and how the residents perceived their competencies in representative clinical situations. A series of clinical vignettes was used to elicit the residents' responses in this last category of the questionnaire.

The demographic data requested were the location of the residency, the year the residency was completed, the graduate's board certification status in internal medicine and pediatrics, whether the graduate's current clinical practice was in primary care, a subspecialty, or both, and whether it was solo or group- or hospital-based, the hours the graduate was involved with patients of different age groups, and his or her various professional activities and fellowship training. Residency format data requested (for both internal medicine and pediatrics components) were the total number of months the resident spent in each department as well as months spent in various rotations (inpatient wards; adult, pediatric, and neonatal intensive care; emergency room; subspecialty electives; primary care and ambulatory care) and whether the graduate felt the time he or she had been involved in each rotation had been too much, too little, or appropriate. A chi-square analysis compared the graduates who reported either too great or too little time spent in a rotation, in order to evaluate trends for each rotation. A Spearman rank correlation coefficient analysis was made to determine the correlation between the total time spent in a rotation and the perceived appropriateness of the duration of involvement in that rotation.

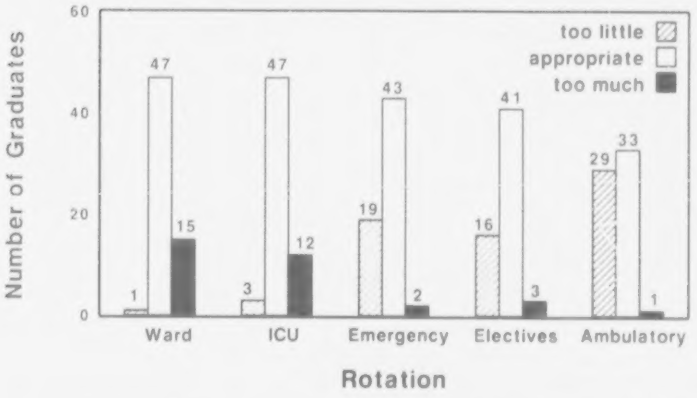


Figure 1. **Clinical rotations in internal medicine: duration and perceived appropriateness of length reported in 1987 by 71 of the 112 graduates of combined internal medicine–pediatrics residency programs.**

The questionnaire listed various subspecialties within each specialty; the graduates were asked to note which of these subspecialties were offered as electives by their residency program, and whether they had completed a particular elective during training. They were asked to rank (from their current perspective) the importance of taking each of the electives, using a Likert-type scale (2 = very important, 1 = somewhat important, and 0 = of very little importance). A contingency table was constructed for comparing the perceived importance scores with whether the rotation was taken. A chi-square analysis was performed, and a contingency coefficient was calculated.

The graduates were asked to note to what degree the internal medicine or the pediatrics housestaff accepted them (very much, somewhat, very little), and asked to outline their continuity clinic experience during the residency. To address whether the combined residency met their expectations, they were asked to report whether they had anticipated, at the beginning of their residency, their current professional activities or practice, or how these activities were different from their expectations. They also were asked to list whether they would again choose a combined residency program. Finally, they were asked to list, in order of priority, recommended changes for med-ped residencies.

Results

Of the 88 graduates from 16 programs mailed surveys directly, 65 (74%) responded. Of the 24 graduates who were to have questionnaires mailed through six other residency programs, six (25%) returned them. The majority of the graduates (and the responding graduates) came from four programs, with by far the largest number of the latter (27) coming from the last year surveyed, 1986. Not all the graduates who filled out the questionniare responded to every question, so that the number of graduates' responses for a particular question may not equal 71.

Internal Medicine Component

During residency training, the responding graduates spent a mean of 24.0 months in internal medicine rotations, with a range of 12 to 30 months. As noted in Figure 1, they had a mean of 10.0 months in inpatient ward rotations, 6.7 months on elective, 3.7 months in intensive care (ICU), 2.6 months in ambulatory rotations, and 1.8 months in emergency room (ER) rotations. The graduates were more likely to feel there was too much, rather than too little, time on ward and ICU rotations ($p < .001$ and $p < .05$, respectively) and too little time in ER rotations, electives and ambulatory rotations (chi square, $p < .001$, $p < .01$ and $p < .001$, respectively). The graduates noted the amounts of time they spent on various rotations, and their perceptions of the appropriateness of the length of involvement. Except for ward and ambulatory rotations, the graduates who had longer involvements with a given rotation felt a shorter duration would have been more appropriate.

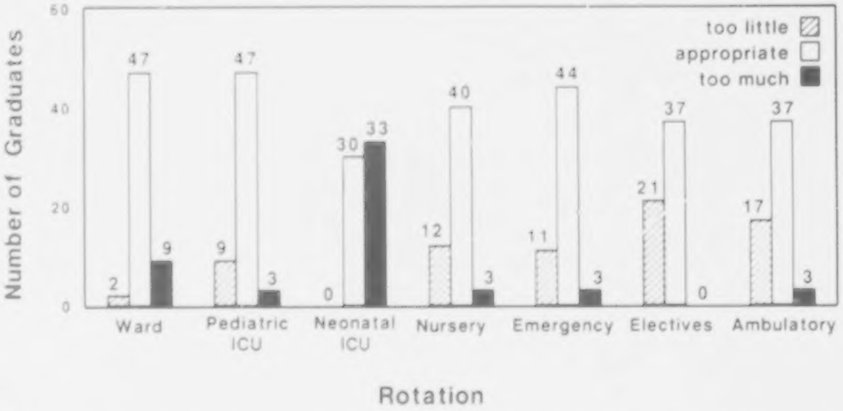


Figure 2. **Clinical rotations in pediatrics: duration and perceived appropriateness of length reported in 1987 by 71 of the 112 graduates of combined internal medicine–pediatrics residency programs.**

The responding graduates ranked 15 adult subspecialty rotations by perceived importance and by whether they had taken those rotations. The cardiology, dermatology, and pulmonary rotations (in that order) were felt to be the three most important. Of the rotations perceived to be among the more important, dermatology and gynecology are nontraditional internal medicine subspecialties. There was often a significant correlation between the graduates' perceiving a rotation as important and having taken that rotation. However, for dermatology, neurology, rheumatology, ophthalmology, and orthopedics, whether it was taken (or not) did not predict whether it was felt to be important. The graduates also noted rotations considered somewhat or very important that their training program had not offered during residency. These included gynecology (12 respondents) orthopedics (10), psychiatry (7), ophthalmology (6), and allergy-immunology (5).

Pediatrics Component

During training, the responding graduates spent a mean of 23.9 months in pediatrics, with a range of 18 to 30 months. Specifically, they spent a mean of 7.9 months in ward rotations, 5.4 months in the ICU (1.7 pediatrics and 3.7 in neonatal), 1.2 months in the normal newborn nursery, 4.7 months in electives, and 3.5 months in ambulatory care. The graduates tended to feel they had spent too much, rather than too little, time on wards and in the neonatal ICU ($p < .05$, $p < .001$), and too little time in the newborn nursery ($p < .05$), in the ER ($p < .05$), on electives ($p < .001$), and in ambulatory care ($p < .01$) (Figure 2). Significant correlations ($p < .05$) between the length of time spent and the perceived appropriateness of a rotation were present for all rotations except the ward and emergency rotations. Each graduate ranked 12 pediatrics subspecialty rotations by perceived importance, and noted whether he or she had taken a given rotation. Infectious disease, cardiology, and adolescent medicine were felt to be the most important subspecialties. For several rotations, there was a poor correlation between rotations thought to be important and the graduate's having taken them, including pediatric cardiology, endocrinology, hematology-oncology, surgery, and orthopedics. Several of the graduates did not have an opportunity to take selected pediatrics rotations that they had felt to be somewhat or very important during their residencies, including orthopedics (11 graduates), surgery (7), adolescent medicine (5), and gastroenterology (4).

Other Curriculum Issues

The respondents felt very accepted by the traditional house officers during training. Sixty-three reported feeling very accepted by the internal medicine residents (eight reported moderate acceptance, none felt not accepted). Similarly, 60 felt very accepted during training by the pediatrics residents (11 reported moderate acceptance, none felt not accepted).

The majority of the graduates (71%) had separate half-day internal medicine and pediatrics continuity clinic experiences weekly. Some (28%) had attended a weekly combined med-ped clinic, and a smaller number had a continuity clinic once a week that alternated between internal medicine and pediatrics. The graduates reported a mean of 1.8 continuity clinic sessions per week.

Sixty-four of the participants (93% of those who responded to this question) reported they would again choose a combined med-ped residency, and all but one felt the residency met their expectations. When requested to list and rank recommended changes, the graduates most commonly recommended an increase in the amount of time involving outpatient subspecialty rotations, specifically in gynecology, orthopedics, and adolescent medicine (Table 1). However, the greatest number of graduates ranked as most important that there be an increased amount of time for ambulatory care experiences within internal medicine and pediatrics. Graduates frequently recommended a combined med-ped continuity clinic (31% of those without a combined clinic recommended one). Decreasing the time spent in intensive care (especially the intensive care nursery), addressing administrative concerns, and increasing the preparation for nonmedical aspects of care (including family therapy or financial aspects of health care) were other significant responses of the graduates.

Table 1

Changes in Training Recommended by Graduates of All U.S. Combined Internal Medicine–Pediatrics Residency Programs, 1987*

Recommendation	Number of Graduates Who Listed the Recommendation	Number of Graduates Who Ranked the Recommendation as Most Important
Increase time in outpatient subspecialty rotation†	23	10
Decrease time in intensive care	18	10
Increase/intensify ambulatory care experience	17	13
Offer combined internal medicine–pediatrics continuity clinic	15	8
Establish more defined curriculum, have a designated program director, greater identification as a program, regular seminars, and more recreational activities	14	6
Address administrative issues, have greater flexibility or cooperation between internal medicine and pediatrics programs	9	2
Make available career counseling and information about financial aspects of practice, psychosocial aspects of care, and methods of family therapy	5	—

*The 112 graduates of all programs of this type were surveyed in 1987 regarding information about their residencies, their current practices, and recommendations for changes in the training of residents in such combined programs. A total of 71 graduates responded to the questionnaire.

†Of the number of graduates who listed this recommendation, six were referring to gynecology rotations, six to orthopedics, four to adolescent medicine, and the other seven to rotations in various other specialties.

Current Activities

Fifty-five of the graduates had already received certification in internal medicine, 60 had passed the written portion of the pediatrics examination, and 34 had passed the oral portion (37 of the respondents had graduated too recently to be eligible for the oral portion). Twenty-three of the graduates had pursued fellowship training and one had completed a second residency in preventive medicine. The fellowships included three in infectious diseases and two each in adolescent medicine, gastroenterology, cardiology, general pediatrics, pediatric pulmonology, neonatology, and the Robert Wood Johnson Clinical Scholars program. There was one fellow each in rheumatology, oncology, hematology, pediatric hematology-oncology, general medicine, and behavioral pediatrics. Twelve of the fellowships combined training in treating both child and adult populations.

The respondents reported a mean of 42.9 hours per week of clinical work; and 83% were seeing patients in both child and adult age groups. The mean distribution of hours spent in the different age groups was 4.9 for newborns, 5.1 for toddlers, 4.9 for school-aged children, 4.9 for teenagers, 7.0 for young adults, 9.3 for the middle-aged, and 6.8 for the elderly. Although the graduates spent most of their time in patient care (42.9 hours), they also noted significant time in research (7.0 hours), teaching (5.8 hours), and administration (3.0 hours). Many practiced in a group setting. The majority were involved in primary care only (64%), with an additional 16% in subspecialty care only.

When asked if their current activity were the same as that anticipated from the start of their residency, 39 of the 71 (55%) replied no. Approximately half of those who said no reported that they either had anticipated a career in primary care (16) or had not planned their current career in academic medicine (3). Although a majority of the graduates were able to practice in a combined med-ped setting, 11 reported that they had not been able to; five of these were not practicing pediatrics, three were not practicing internal medicine, and three were practicing both internal medicine and pediatrics but were unable to integrate the two specialties at a single practice setting.

Discussion

This first nationwide survey of graduates of combined medicine and pediatrics residency programs has shown that although these graduates were generally satisfied with their training, they had several specific criticisms of their curricula. During both their internal medicine and their pediatrics training time, the responding graduates would have preferred more ambulatory care and less inpatient experience. There was an especially negative response toward the neonatal intensive care unit. This is not surprising, given the disproportionate amount of time devoted to this area in most pediatrics programs. Indeed, in the mid-1970s, the American Academy of Pediatrics reported that almost half of their surveyed membership thought they had had too much newborn care training during their residencies.[6] Despite recommendations at that point to limit such exposure, a 1984 survey showed little change in the number of months that pediatrics residents devoted to the area of newborn care.[7]

The graduates generally wanted to have had more ambulatory care time, including primary care, emergency care, and subspecialty electives. Those outpatient specialty experiences that were especially desired were gynecology, orthopedics, and adolescent medicine. These findings are in accord with other published studies of graduates in internal medicine residencies, which show that the graduates consistently want more training in common ambulatory care problems and in subspecialty ambulatory care.[8] Practitioners of general medicine who have trained in either traditional or primary care medicine programs feel underprepared in such "nonmedical" disciplines as orthopedics, dermatology, psychosocial problems, and gynecology.[8-12] Adolescent medicine, however, stands out as a part of training that many med-ped graduates desired. Because of their exposure to both children and adults, some med-ped physicians may see practicing adolescent medicine as particularly appropriate, and it is possible that other physicians would see them as the best consultants concerning this crossover age group. Without directed training in the unique morbidities and developmental issues of adolescence, however, a med-ped physician may be more qualified to treat these patients than any other primary care physician.

The desire of the med-ped graduates surveyed in this study to have more ambulatory education is in agreement with the general trend in medical education away from the hospital ward and toward the outpatient setting, an ever-expanding site of patient management.[13] Even graduates of traditional, subspecialty-oriented internal medicine programs desired increased training emphasis in areas most applicable to outpatient care.[10]

One part of a trainee's primary care experience in an ambulatory care setting is the continuity clinic, an experience that residents, at least in internal medicine, have not always found to be of high quality.[9] In med-ped programs, many different combinations of weekly or semiweekly internal medicine, pediatrics, or combined practices exist. Those who did not experience a combined med-ped continuity clinic often recommended a change to this arrangement that would allow residents to train to see patients of all ages in a single session. Doing this is important because establishing a practice that combines all age groups in a goal anticipated by almost all med-ped residents.

While there was usually a correlation between graduates' having taken a particular subspecialty and their perceiving that the rotation in that subspecialty was important, five medicine and five pediatrics subspecialties were an exception to this trend. In eight of these ten cases, the lack of correlation was a result of the fact that even those who did not take the particular rotation deemed it important. The implication is that med-ped graduates, like internal medicine graduates, wish for exposure to more subspecialty areas without sacrificing the ones they have.[8] As Linn and colleagues[8] note, "This phenomenon of wanting more without an expressed willingness to make cuts elsewhere proves a continuing dilemma to medical educators who try to be responsive to the increasing and often changing needs of residents." In the case of med-ped programs, one way of offering more subspecialty experiences without consuming more valuable time is to offer combined subspecialty rotations for fields such an endocrinology, rheumatology, infectious disease, gastroenterology, and nephrology.

From these data one can first conclude that med-ped training is flexible, allowing either academic or community practice and either general or subspecialty care. Of the 71 graduates who responded, 23 had already pursued fellowship training, and more than half of these fellowships were for a combined med-ped population of patients. The ability to practice a subspecialty in both internal medicine and pediatrics areas is well established for those trained in med-ped. Further, the graduates studied have generally been successful in maintaining practices that involve both adult and pre-adult patients; 83% of the graduates saw patients of all ages. The versatility of training is further supported by board certification performance. Passing rates for med-ped graduates taking the pediatrics and internal medicine board exams are at least as high as those of their counterparts in each of these specialties.[5] Although this may not be an accurate indicator of one's ability to practice both internal medicine and pediatrics, it does establish comparable performance between combined-specialty and non-combined-specialty residents in this one measurement of cognitive skill.

What is implied in the diversity of graduates' careers is that the different med-ped residency programs have different strengths. Some programs are more oriented toward primary care and some more oriented toward their graduates' pursuing subspecialties or other academic pathways. This has been reflected in the opinion of directors of such programs.[14] It may be important, therefore, for each med-ped program to develop and explicitly state its focus. This would not only assist each program to define and nurture its own strengths, but would also help prospective residents to choose the programs that most nearly fit their needs.

There are several possible shortcomings to the present study. By attempting to poll all med-ped graduates as of July 1986 through contact with programs listed in the 1986–1987 *Directory of Residency Training Programs*, the investigators probably omitted some because there have been graduates either of programs that no longer exist or of ad hoc programs not recognized by the NRMP. These potential omissions probably represent a small portion of the total population of graduates. Moreover, from the standpoint of introducing curricular innovation into med-ped training, it is more valuable to solicit feedback from graduates of currently functioning programs, as has been done. The difference in response rates between those graduates mailed questionnaires directly and those mailed questionnaires indirectly through program directors is probably because some directors of their staffs did not forward the questionnaire. However, there does not appear to be any systematic bias introduced into the results because of this problem.

The interpretation of results form this survey must also be tempered by the knowledge that med-ped programs are rapidly expanding and thus information will be changing rapidly. Although the oldest programs graduated residents more than 15 years ago, the majority of graduates who responded (38 of 71) had finished training in 1985 and 1986; 27 finished in 1986 alone. Therefore, such data as board certification rates, the proportion of graduates ultimately pursuing fellowships, and practice patterns, all of which may take several years to stabilize, are by necessity preliminary. The proportion of residents pursuing primary care careers, for instance, may change over time. The program with the greatest number of graduates to date (the University of North Carolina at Chapel Hill) has produced proportionally more academicians and subspecialists, whereas the second largest (the University of Rochester School of Medicine and Dentistry) has a larger proportion entering primary care.[1] Most of the newer programs tend to emphasize primary care training, and over time will probably account for a growing share of graduates. A disturbing trend emerging at some centers, as found by Siegel and colleagues,[14] is the establishing of a med-ped program in order to strengthen a relatively weak internal medicine or pediatrics program. If this trend continues, it may lead to less success for med-ped graduates in board certification scores and in the ability to sustain combined practices and obtain quality fellowships.

Further research is needed to address these questions now that the rapid growth of numbers of graduates has begun to level off.[2] Although graduates are a valuable source of data, faculty and current trainees may offer complementary information regarding curricular issues. In addition, a major question only hinted at by the present survey is the recent rapid growth of med-ped programs. Med-ped program directors (a small minority of whom were trained in med-ped) most commonly view this combined training as an alternative primary care pathway to family medicine.[14] It is unclear how graduates, current residents, or fourth-year medical students view this issue.

In conclusion, med-ped training has undergone tremendous expansion, approaching, in the mid 1980s, the number of applications to residency programs in primary care internal medicine. Although graduates of med-ped programs have generally been satisfied with their training and successful in their careers, they have certain criticisms of their residency curricula. Their most important recommendation are (1) to limit the required amount of newborn care (especially neonatal ICU), perhaps to one-sixth of all pediatrics time; (2) to emphasize opportunities for ambulatory care training and offer nontraditional outpatient subspecialty rotations such as orthopedics and gynecology, combined med-ped continuity clinics, and adolescent medicine training; and (3) to offer combined rotations in those subspecialties that are more outpatient-oriented and less procedure-oriented.

This study was funded in part by Maternal and Child Health Grant #OH964.

The authors acknowledge Robert A. Hoekelman, M.D., and Joseph L. Rauh, M.D, for their careful review and suggestions, and Lynn Hanrahan for clerical assistance.

References

1. Greganti, M. A., and Schuster, B. L. Two Combined Residency Programs in Internal Medicine and Pediatrics. *J. Med. Educ.* **61**(1986):883.
2. *NRMP Data.* Evanston, Illinois: National Resident Matching Program, 1988.
3. Graettinger, J. S. Internal Medicine in the National Resident Matching Program 1987: The Ides of March. *Ann. Intern. Med.* **108**(1988):101-115.
4. Ferrari, N. D. Hybridization of the Primary Care Disciplines. *JAMA* **256**(1986): 2345.
5. Shumway, J. M., and Ferrari, N. D. Survey of Medicine/Pediatrics Residency Training Programs. *J. Gen. Intern. Med.* **2**(1987): 377-380.
6. The Future of Pediatric Education: A Report by the Task Force on Pediatric Education. Evanston, Illinois: American Academy of Pediatrics, 1978.
7. Weinberger, H. L., and Oski, F. A. "A Survey of Pediatric Resident Training Programs 5 Years after the Task Force Report". *Pediatrics* **74**(1984):523-526.
8. Linn, L. S., Brook, R. H., Clark, V. A., Fink, A., and Kosecoff, J. Evaluation of Ambulatory Care Training by Graduates of Internal Medicine Residencies. *J. Med. Educ.* **61**(1986):293-302.
9. Kantor, S. M., and Griner, P. F. Educational Needs in General Internal Medicine as Perceived by Prior Residents. *J. Med. Educ.* **56**(1981):748-756.
10. Kern, D. C., Parrino, T. A., and Korst, D. R. The Lasting Value of Clinical Skills. *JAMA* **254**(1985):70-76.
11. Mitchell, W. D. and Thompson, T. L. Attitudes of Internists in Practice toward Their Ambulatory Care Education. *J. Med. Educ.* **60**(1985):415-417.
12. McPhee, S. J., Mitchell, T. F., Schroeder, S. A., Perez-Stable, E. J. and Bindman, A. B. Training in a Primary Care Internal Medicine Residency Program. *JAMA* **258**(1987):1491-1495.
13. Schroeder, S. A., Showstack, J. A., and Gerbert, B. Residency Training in Internal Medicine: Time for a Change? *Ann. Intern. Med.* **104**(1986):554-561.
14. Siegel, D. M., Parker, R. M. Gillman, M. W., and Biro, F. M. Demographic Features and Attitudes of Program Directors of Combined Internal Medicine and Pediatric Residencies. AJDC **142**(1988):1104-1108.

AAMC DATA REPORT

Women on Faculties of U.S. Medical Schools, 1978–1989

In the 11 years between 1978 and 1989, the number of women faculty members in U.S. medical schools increased 76%, compared with a 25% increase of men faculty (Table 1). But women faculty still comprise only 20% of the total full-time faculty, up from 15% in 1978 and 18% in 1985. The growth has been greater among women with M.D. and M.D.-Ph.D. degrees than among women with Ph.D. or other doctorate degrees in health (such as public health, dental medicine and surgery, and veterinary medicine). There has been a larger percentage growth in men with Ph.D. degrees than in those with M.D. degrees. The numbers and proportions of both men and women faculty with "other" (predominately master's) degrees have declined.

The percentage of minority women on faculties changed little over the 11-year period, remaining close to the proportions shown in Figure 1 (page 278). However, the rate of increase in the number of minority women was greater than that for white women faculty, and women also tended to be a larger proportion of the minority faculty than the white faculty (Figure 2).

Although the numbers of women in all faculty ranks grew, the increase was greatest among women assistant professors (140%), and 49% of women faculty members now hold this rank (Table 2). Only 29% were at the associate professor or professor levels, compared with 57% of men. Although proportionately more women than men graduates of medical school become full-time faculty members,[1,2] the fact that only about a third of the graduates are women curtails the total number of women entering faculties. Indeed, since 1985 the proportion of women at the associate professor level has not grown, and that of those at the level of professor has dropped slightly.

As Figure 3 shows, the numbers of women in all medical school departments increased significantly. The percentage increases were far from uniform; and in pediatrics, which has a far higher representation of women than any other specialty, the percentage of women actually declined slightly.

While these data show that the number of women on U.S. medical school faculties continues to grow, there remains wide variation in their proportions across academic departments. Better understanding of the lag in progression of women to associate professor and professor ranks probably will not emerge until schools begin to examine their promotion requirements and the guidance and career development of their faculty women. (All data for this report are from the AAMC's Faculty Roster, which contains information on most full-time faculty at U.S. medical schools.)

Brooke E. Whiting, Ph.D.
Janet Bickel, M.A.

Dr. Whiting is director, Faculty Roster, and Ms. Bickel is director for women's programs, Division of Institutional Planning and Development, Association of American Medical Colleges, Washington, D.C.

References

1. Jolly, P., and Larson, T. A. Women Physicians on U.S. Medical School Faculties. *J. Med. Educ.* **50**(1975):825–828.
2. Jolly, P. Women Physicians on U.S. Medical School Faculties. *J. Med. Educ.* **56**(1981):151–153.

Table 1

Distribution of U.S. Medical School Faculty by Degrees and Gender in 1978 and 1989

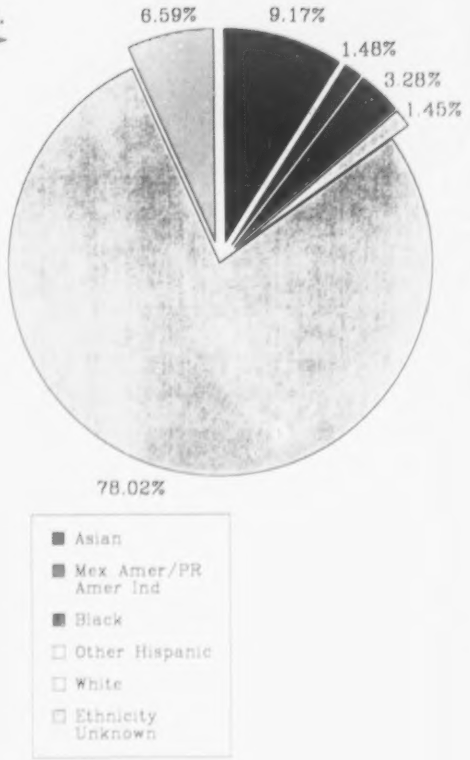
	1978				1989			
	Men		Women		Men		Women	
	No.	%	No.	%	No.	%	No.	%
M.D. and M.D.-Ph.D.	27,708	69	3,174	44	34,894	69	6,859	54
Ph.D., and other health doctorates	9,696	24	1,916	27	13,530	27	3,897	31
Other degrees	2,583	6	2,063	29	1,607	3	1,857	15
Total	39,987	100	7,153	100	50,031	100	12,613	100
Women as % of total				15				20

Table 2

Numbers and Percentages of Women Faculty by Rank

	1978		1985		1989	
Rank	No.	%	No.	%	No.	%
Professor	518	7	892	10	1,170	9
Associate professor	1,123	16	1,867	20	2,455	20
Assistant professor	2,597	36	4,402	47	6,222	49
Instructor	1,542	22	1,746	19	2,244	18
Other	469	7	381	4	507	4
Missing	904	13	—	—	15	0
Total	7,153	100	9,288	100	12,613	100

Figure 1. **Race-ethnicity of women faculty in U.S. medical schools, 1989.**



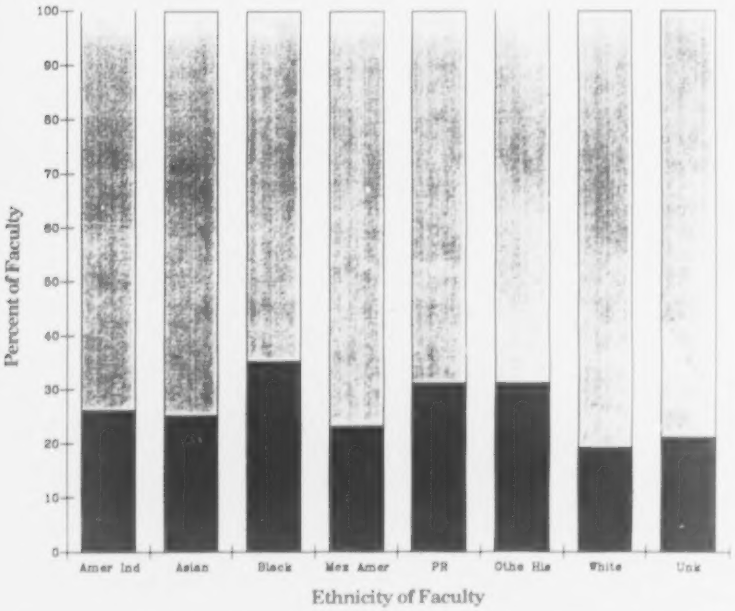


Figure 2. **Distribution of medical school faculty by race-ethnicity and gender, 1989. Men, light shading, women, dark shading.**

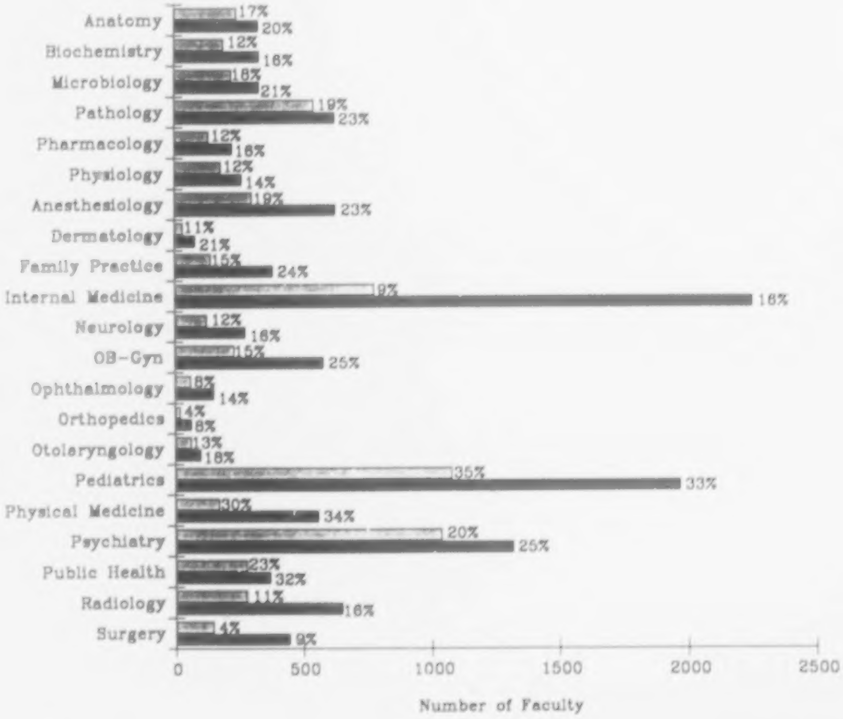


Figure 3. **Number of women and percentage of total faculty, by department, in U.S. medical schools, 1978 (light shading) and 1989 (dark shading).**

A3273-8

MAY 1990

ACADEMIC MEDICINE

Recruiting Black Medical Students

Assessment

- Use of Standardized Patients
- Measuring Program Outcomes
- Clinical Skills at Four Schools
- Medicine Clerks

Implications of Primary Health Care

Medical Informatics at Dartmouth

Problems of Late Hospitalization

TIME Patient Simulation Model

Integrating Clinical Data

Family Medicine Training

VOLUME 65 • NUMBER FIVE

MAY 1990
VOLUME 65
NUMBER FIVE

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Invited Articles

Recruiting Black Medical Students: A Decade of Effort, *by Robert E. Taylor, Jr., Ph.D., James C. Hunt, M.D., and Patra B. Temple, Ed.D.* 279

Results of a Survey on the Use of Standardized Patients to Teach and Evaluate Clinical Skills, *by Paula L. Stillman, M.D., Mary Beth Regan, Ed.D., Mary Philbin, and Heather-Lyn Haley* 288

The Measurement of Outcomes in the Assessment of Educational Program Effectiveness, *by Donald G. Kassebaum, M.D.* 293

Ideas for Medical Education

Using Classic and Contemporary Visual Images in Clinical Teaching, *by Janine C. Edwards, Ph.D.* 297

Medical Informatics

Medical Informatics for the Other Ninety Percent: The Dartmouth Experience, *by J. Robert Beck, M.D., Edward K. Shultz, M.D., M.S., and Brian R. Edwards, Ph.D., MS* 298

International Medical Education

Primary Health Care: Implications for the Medical Profession and Education, *by Snehendu B. Kar, Dr.P.H., M. Sc.* 301

Editorial

On Standardized Patients and Clinical Skills Assessment, *by Donald G. Kassebaum, M.D.* 307

Letters to the Editor

Residents' Hours and Supervision, *by Michael D. Cusimano, M.D., and William S. Tucker, M.D.* 308

Specialty Choice, *by Janet Bickel, M.A.* 308

Predicting Matriculation, *by Marc S. Nelson, M.D.* 308

From the President 310

AAMC Focus 311

National Policy Perspectives 312

Research Reports
ARTICLES

The Problem of Late Hospitalization: A Quality and Cost Issue, *by Joseph* S. *Gonnella,* M.D., *Daniel* Z. *Louis,* M.S., *Carter Zeleznik,* Ph.D., *and Barbara* J. *Turner,* M.D. 314

An Assessment of the Clinical Skills of Fourth-Year Students at Four New England Medical Schools, *by Paula* L. *Stillman,* M.D., *Mary Beth Regan,* Ed.D., *David* B. *Swanson,* Ph.D., *Susan Case,* Ph.D., *John McCahan,* M.D., *Joel Feinblatt,* Ph.D., *Stephen* R. *Smith,* M.D., *Janice* W. *Willms,* M.D., *and* D. *Virginia Nelson* 320

A Field Test of the TIME Patient Simulation Model, *by William* G. *Harless,* Ph.D., *Robert* C. *Duncan,* Ph.D., *Marcia* A. *Zier, William* R. *Ayers,* M.D., *Jerome* R. *Berman,* M.D., *and Henry* S. *Pohl,* M.D. 327

A Comparison of Two Methods for Facilitating Clinical Data Integration by Medical Students, *by* I. *Jon Russell,* M.D., Ph.D., *William* D. *Hendricson,* M.S., *Gary* D. *Harris,* M.D., *and Denise* V. *Gobert* 333

Evaluation of Students in Medicine Clerkships, *by Gregory* J. *Magarian,* M.D., *and Dennis* J. *Mazur,* M.D., Ph.D. 341

A New Faculty Training Model for Countries Developing Academic Family Medicine Programs, *by Douglas* H. *Johnson,* M.D., *and Alfred Zammit-Montebello,* M.Ed. 346

SHORT PAPER

The Availability of Fellowship Training for Foreign Family Physicians, *by Michael* D. *Cohen,* M.B., Ch.B., *and Joel* H. *Merenstein,* M.D. 348

Publishing Services: **Hanley & Belfus, Inc.**
210 South 13th Street, Philadelphia, PA 19107

MAY 1990
VOLUME 65
NUMBER FIVE

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
Edward J. Stemmler, M.D.
Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
Arizona Health Sciences Center

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
Georgetown University Hospital
Philip Anderson, M.D.
University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
Harvard Medical School
Eta Berner, Ed.D.
University of Alabama School of Medicine
Ruth Bulger, Ph.D.
National Academy of Sciences
Vincent A. Fulginiti, M.D.
Tulane University School of Medicine
Nancy Gary, M.D.
Albany Medical College
Beatrix Hamburg, M.D.
Mt. Sinai Hospital
Murray M. Kappelman, M.D.
The University of Maryland School of Medicine
Ralph W. Muller
University of Chicago Hospitals
Herbert Pardes, M.D.
College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
University of Alabama School of Medicine
W. Vickery Stoughton
Toronto General Hospital
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
David H. Cohen, Ph.D.
Northwestern University

Chairman-Elect
William T. Butler, M.D.
Baylor College of Medicine

Immediate Past Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center School of Medicine

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Joe Dan Coulter, Ph.D.
University of Iowa College of Medicine

Chairman, Council of Teaching Hospitals
Raymond G. Schultze, M.D.
University of California, Los Angeles Medical Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.
Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036; (202) 828-0416.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

5. *Minority Students in Medical Education: Facts and Figures IV.* Washington, D.C.: Association of American Medical Colleges, 1988.
6. *Statistical Abstract of the United States: 1987.* Washington, D.C.: U.S. Bureau of the Census, 1987.
7. Personal communication, June 13, 1988. Tennessee Department of Vital Statistics, Nashville, Tennessee.
8. Personal communication, September 6, 1989. Tennessee Department of Health and Environment, Division of Information Resources, Nashville, Tennessee.
9. Personal communication, June 21, 1988. Association of American Medical Colleges, Student and Applicant Information Management System Data Base.
10. Etzel, S. I. Medical Education in the United States 1978-1979. *JAMA* **243**(1980):841-866.
11. *Desegregation Monitoring Report.* Nashville, Tennessee: Tennessee Higher Education Commission, April 1989.
12. Personal communication, June 15, 1988. Association of American Medical Colleges, Faculty Roster System Data Base.
13. Minority Admission Score Card. *The New Physician.* **35**(December 1986): 20-21.

PAULA L. STILLMAN, M.D., MARY BETH REGAN, Ed.D., MARY PHILBIN, and HEATHER-LYN HALEY

Results of a Survey on the Use of Standardized Patients to Teach and Evaluate Clinical Skills

Abstract—In 1989, a survey was sent to each U.S. and Canadian medical school requesting information about how standardized patients are used for teaching and evaluating clinical skills, and 95% of the schools responded. Although there was widespread use of standardized patients throughout the curricula, the role and training of these patients varied markedly within a given school as well as across schools. One outcome of this survey is the development of a network to share resources, protocols, and training material to enhance the development of this educational strategy. *Acad. Med.* 65(1990):288-292.

Standardized patients (SPs) have been used throughout medical school curricula for at least 25 years[1-8] to teach and evaluate clinical skills. SPs, often called simulated patients, patient instructors, or programmed patients, are lay persons trained to replicate a clinical encounter consistently and realistically. They may have no specific medical background, may be actors, or may be actual patients with or without real physical abnormalities. The SP may only simulate a patient or, with additional training, may record aspects of the student's performance and evaluate the student's skills on specially designed forms. With further training, the SP may function as a teacher and at the end of the encounter provide the student with individualized and constructive feedback. According to the results of the 1989 LCME Part II Medical Education Questionnaire, SPs are being used at an increasing number of schools.[9] Although a majority of U.S. medical schools indicate that they use standardized patients in the curriculum, the extent to which and purposes for which they are used are not generally known. Many medical educators believe that SPs are a valuable tool in a variety of clinical exercises to provide students with uniform teaching and evaluation of basic clinical skills and to ensure consistent exposure to common ambulatory problems. One of the purposes of this survey was to document the extent of this use and to assess the interest of medical schools in establishing a network that would facilitate the development or expansion of SP programs.

Method

In July 1989, a survey on the use of standardized patients for teaching and evaluating medical students and residents was mailed to 142 curriculum deans at U.S. and Canadian medical schools. After five weeks, a second copy of the survey was mailed to all that had not replied. A total of 136 (95%) responded to the survey. (The detailed results are available from the first author.) The survey questionnaire requested detailed documentation on the use of SPs throughout the curriculum. Data were collected to answer these questions: (1) How are SPs used at various institutions to teach and evaluate through the curriculum? (2) What are common problem areas? (3) What are some of the characteristics of successful programs? (4) What cases, evaluation instruments, etc., are there that can be

---

Dr. Stillman is curriculum dean, Dr. Regan is an education specialist, Ms. Philbin is coordinator of the standardized patient program, and Ms. Haley is a research assistant, all at the University of Massachusetts Medical Center, Worcester.

Correspondence and requests for reprints should be addressed to Dr. Stillman, Office of the Curriculum Dean, University of Massachusetts Medical Center, 55 Lake Avenue North, Worcester, MA 01655.

Related articles about standardized patients appear on pages 307, 320, and 327.

shared? and (5) Is there interest enough for a resource network to be established among academic institutions?

Results

The results are summarized here in terms of 16 topics. Although the survey results are from 1989, the findings can be considered current because SP programs tend to be stable and long-term.

Use of SPs. Standardized patients are used in some way at 94 (70%) of the 136 schools that responded to the survey. The survey results presented in Figure 1 illustrate the numbers of schools using SPs across the curriculum for teaching and evaluation. The numbers on the graphs are higher than the total number of schools listed in the text because SPs are often used in more than one curriculum year.

Breast and Pelvic Examination. The most popular use of SPs is to train students in breast and pelvic examination. This type of program now exists at 84 medical schools. It mainly occurs in the second and third years of the curriculum, with SPs used primarily to teach. Each teaching session lasts 10 to 62 minutes. The schools that use SPs to evaluate students or residents allot 5 to 25 minutes for that portion of the encounter. Figure 1A summarizes these uses of SPs.

Male Genitourinary (GU) Examination. SPs are used for the male GU exam at 45 schools, mostly in the second year and primarily for teaching purposes. The length of encounter is between 10 and 40 minutes, and evaluation of the encounter usually 5 to 17 minutes. Figure 1B details the use of SPs for this examination.

Interviewing Skills. Sixty-two schools use SPs for teaching or evaluating interviewing skills. The category "interviewing skills" was defined on the questionnaire as the process by which information is obtained from or shared with the patient. The length of the SP-student encounter is between 17 and 29 minutes, with an additional 6 to 25 minutes being spent on evaluation. A breakdown of SP use by year of training, as well as whether the SP is used for teaching or evaluating the interviewing process, can be seen in Figure 1C.

History Content. Fifty-seven schools use their standardized patients to help the students learn to take a relevant medical history. History content refers to the actual medical history data obtained and does not take into consideration interviewing or interpersonal skills, although overlap certainly exists. The most popular history content areas are the adult, geriatric, and psychiatric histories. The SP-student encounter varies between 14 and 37 minutes, with an additional 7 to 19 minutes for evaluation. The use of SPs for teaching and evaluating history content is illustrated in Figure 1D.

Patient Education and Counselling. Thirty-nine schools use SPs to teach or evaluate counselling skills and patient education. This is defined as the delivery to the patient of information such as the need for behavior modification or lifestyle changes, counselling, or delivering bad news to a patient or the family. The length of the encounter varies from 17 to 32 minutes, with an evaluation time between 10 and 25 minutes. The pattern of usage, with the primary focus on second-year students and residents, can be seen in Figure 1E.

Complete Physical Examination. SPs are used to teach or evaluate the complete physical examination at 33 of the responding schools. The encounter lasts between 20 and 56 minutes, with an additional 5 to 44 minutes spent on evaluation. SPs are used most often in the second year of the curriculum for both teaching and evaluating the complete physical examination, as can be seen in Figure 1F. This aspect of the program is also used for evaluating residents' physical examination skills.

Segments of Physical Examination. Forty-eight schools reported using SPs for teaching and/or evaluating segments of the physical examination. The encounter between the student and the standardized patient lasts between 11 and 30 minutes, with 5 to 26 minutes set aside for evaluation. Figure 1G details the use of SPs in this area, showing that they are used primarily for second- and third-year students.

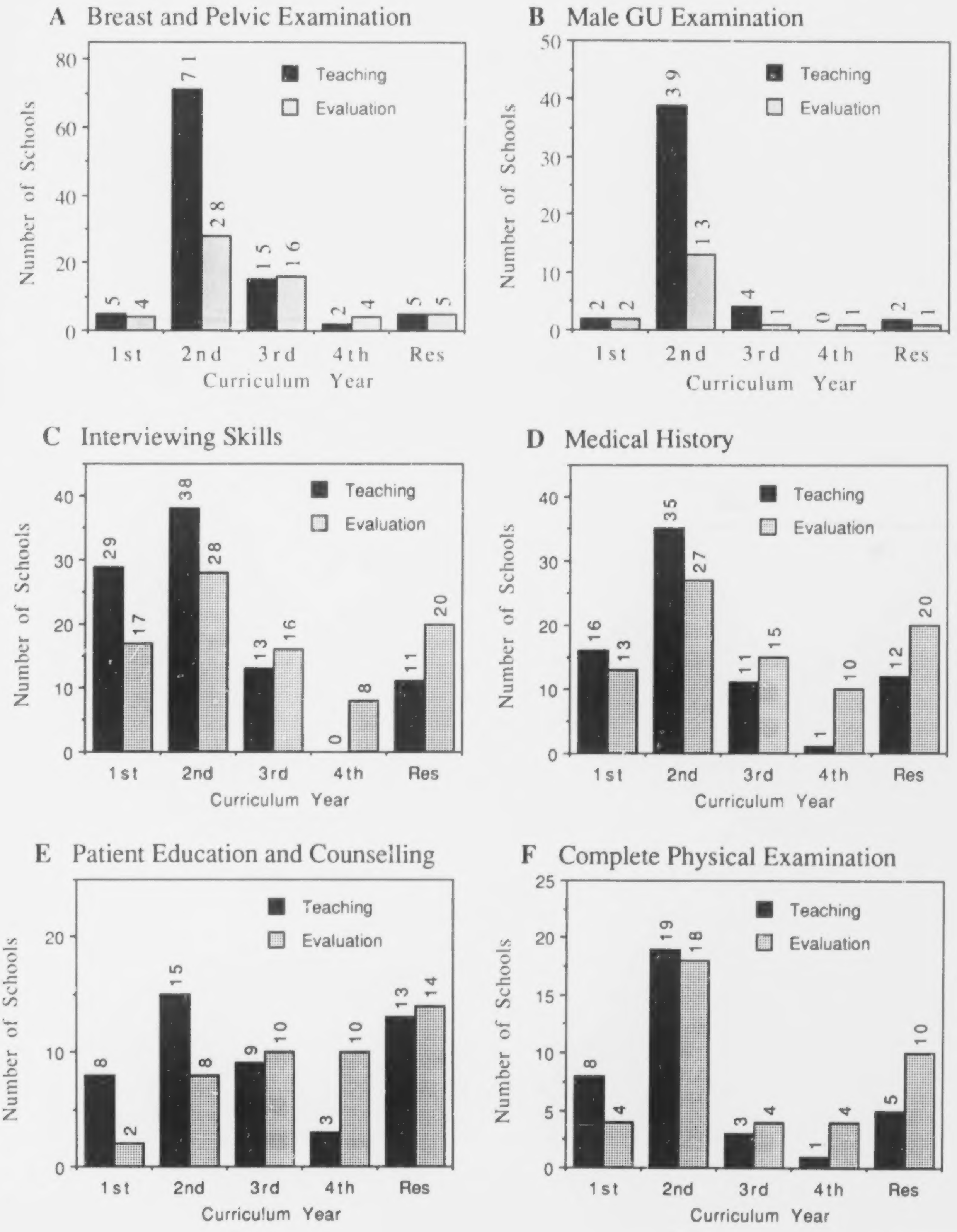
Focused Encounter. Forty-six of the schools surveyed use SPs for focused encounters (brief histories and physical examinations). They are most frequently used for both teaching and evaluation of second-year students. The emphasis shifts to evaluation for third- and fourth-year students and residents, as shown in Figure 1H. The length of encounter is between 14 and 38 minutes, with 9 to 22 minutes spent on evaluation.

Checklist. A majority of schools (64) that use standardized patients have checklists to document students' performance. Depending on the type of encounter, the checklist may be completed by a faculty member, the SP, or both (see Table 1).

Feedback. Feedback is given to students who interact with standardized patients at 87 of the 94 schools that have such a program in place. Ninety-eight percent of the schools provide oral feedback, while 38% give written feedback; 93% give the student feedback immediately following the encounter. Depending on the type of encounter, the feedback may be given by a faculty member, the SP, or both. This information is also summarized in Table 1.

Videotape Use. Forty-three schools videotape encounters with standardized patients. At 33 of these schools, the tapes are used for evaluation of individual students, while 23 schools use them for group teaching and 18 schools, to train new SPs. A majority of taping is done with first-year students and residents. The most popular areas for videotaping are the history content, patient education and counselling, and focused history and physical examinations.

Training of Standardized Patients. Standardized patients are trained by a non-M.D. trainer at 68 of the 94 schools that use SPs, 53 use M.D. trainers, and 27 use the case author. Again, the totals are higher than the number of schools because more than one type of trainer may be used at a single school. The most common training technique is some combination of the individual session (69


A Breast and Pelvic Examination
Teaching
Evaluation
Number of Schools
Curriculum Year
1st 2nd 3rd 4th Res
5 4
71 28
15 16
2 4
5 5
0 20 40 60 80
B Male GU Examination
Teaching
Evaluation
Number of Schools
Curriculum Year
1st 2nd 3rd 4th Res
2 2
39 13
4 1
0 1
2 1
0 10 20 30 40 50
C Interviewing Skills
Teaching
Evaluation
Number of Schools
Curriculum Year
1st 2nd 3rd 4th Res
29 17
38 28
13 16
0 8
11 20
0 10 20 30 40
D Medical History
Teaching
Evaluation
Number of Schools
Curriculum Year
1st 2nd 3rd 4th Res
16 13
35 27
11 15
1 10
12 20
0 10 20 30 40
E Patient Education and Counselling
Teaching
Evaluation
Number of Schools
Curriculum Year
1st 2nd 3rd 4th Res
8 2
15 8
9 10
3 10
13 14
0 10 20
F Complete Physical Examination
Teaching
Evaluation
Number of Schools
Curriculum Year
1st 2nd 3rd 4th Res
8 4
19 18
3 4
1 4
5 10
0 5 10 15 20 25

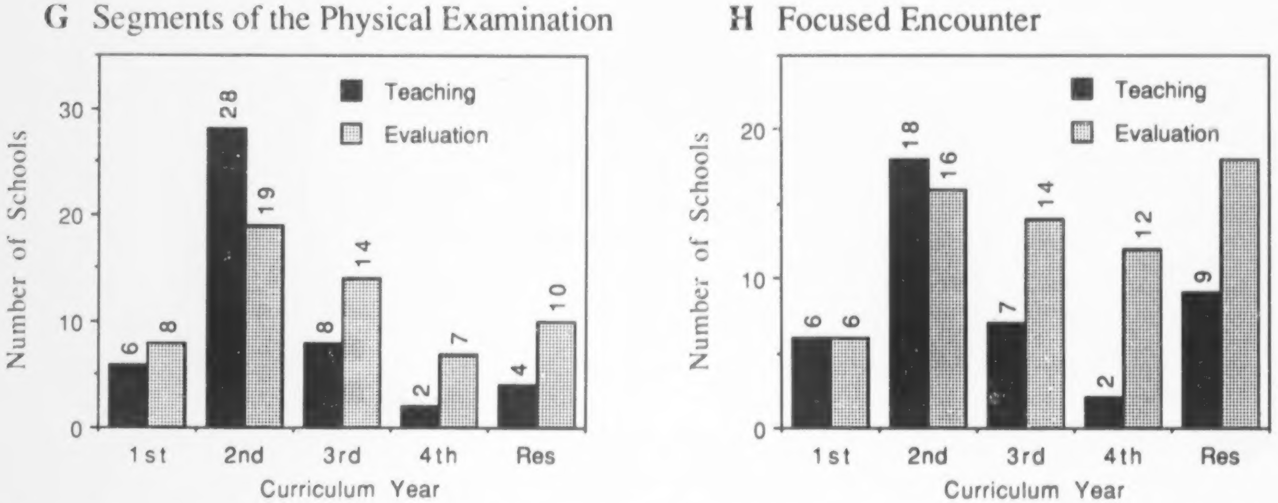


Figure 1 (facing page and above). **Numbers of schools among 136 surveyed in 1989 that used standardized patients for teaching and evaluation in various programs.**

schools) and group session (64 schools), while 41 schools rely on videotapes. Direct observation of other SPs is commonly employed by 42 schools, while 41 schools teach from manuals; only six schools reported training SPs by having them directly observe actual patients.

The average program spends seven hours training the SP to provide a medical history and simulate symptoms or an affective disorder, four hours to teach an SP a focused physical examination, eight hours for the breast and pelvic examination, and six hours for the male GU examination. It takes approximately ten hours to train an SP to evaluate the student's interviewing skills and to provide adequate feedback and teaching. Fifteen hours are required to train SPs to evaluate the student's physical examination skills, including teaching the SPs to provide feedback and teaching in this area.

Administration of Program, Faculty Participation, and Recruitment. Thirty-nine schools with SP programs have a central office that coordinates the program, while 55 do not. The numbers of SPs enrolled in individual programs vary from 2 to 125, with a mean of 27, and the numbers of case protocols range from zero to 400, with a mean of 39. Although one school has used SPs for 25 years, they have been used for a mean of four years at most schools.

Faculty participation is high, with 70 schools reporting that their faculty observe encounters; at 63 schools faculty provide feedback. At 61 schools faculty train the standardized patients, and at 53 schools faculty write cases. Recruitment methods are similar overall, with 53 schools using word of mouth as their main method. Other sources for finding SPs include personal contact, medical and nursing students, senior citizens' groups, campus newspapers and bulletin boards, community support groups such as Alcoholics Anonymous, and drama and acting groups.

Strengths and Weaknesses. Some of the major strengths of the programs surveyed were strong student and faculty support, the importance of "real" patient feedback for students, the opportunity for students to practice counselling skills, and strong case development. Weaknesses mentioned were programs with a limited scope and wasted potential, the need for better defined, stronger evaluation strategies, a lack of organization and time, and inadequate financial support.

Start-up Concerns. For those schools that do not have an SP program, the largest start-up concern is financial, with 28 schools citing it as an obstacle. Fifteen have a problem

Table 1

Numbers of U.S. Medical Schools Using Checklists and Giving Feedback to Students during the Use of Standardized Patients to Teach Eight Skills, 1989

		Checklists Used			Feedback Given		
Skills	Total No.	No.	By Faculty	By SP	No.	By Faculty	By SP
Interviewing skills	62	51	34	37	55	48	39
Medical history content	57	48	31	34	48	43	29
Patient education	39	31	14	23	31	22	18
Complete physical examination	33	27	17	15	23	15	17
Segments of physical examination	48	34	23	17	34	28	18
Breast and pelvic examination	84	35	13	27	68	35	58
Male genitourinary examination	45	19	8	13	34	17	25
Focused encounter	46	34	24	21	31	24	15

with lack of interest from the faculty, while only six schools list administrative lack of interest as a drawback. Fifteen schools lack individuals who have the skills to run the program and train SPs, and nine cite an inability to locate appropriate standardized patients. Another concern is an overall lack of faculty time to devote to developing and sustaining such a program.

Formation of Network. Ninety-one of the 94 schools with an SP program are interested in forming a network. Forty-four schools have cases to share, while 14 offered training videotapes and 15 indicated that they had training manuals. Other materials respondents are willing to share include checklists, objectives, lecture materials, facilities, evaluation forms, and SP handbooks.

Discussion

It is clear from reading the survey results that the term standardized patient implies a variety of entities. Because there is no uniformity in defining the role or training of the SP, it is important that we develop a common classification system so that communication within and across institutions can be enhanced. The term is now used to include the spectrum of real patients without training, simulated patients, and role playing; the function of SPs ranges from simulator, to evaluator, to teacher. There is a pressing need for a hierarchical classification with agreement as to the meanings of specific terms.

We are enthusiastic about the frequent use of SPs to supplement current teaching activities and as part of a diagnostic assessment of students' skills throughout the medical school curricula. Because of the changing nature of the hospitalized patient population and the difficulty in teaching in an ambulatory setting, SPs can be used to ensure that each student has exposure to common ambulatory problems, such as those of patients with newly diagnosed diabetes or anginal chest pain. They can help ensure uniformity in curriculum. They can also be used to teach basic clinical skills such as interviewing and physical examination techniques, and clinicians' time can then be reserved for discussing diagnostic clinical problem solving once basic skills are mastered. SPs are also an excellent addition to problem-based small-group teaching. The cases discussed can come alive by having students interact with an SP who portrays the patient realistically. Several SPs can be trained in the same role, allowing each small group the benefit of this exposure.

Some types of SP cases seem more appropriate for teaching, while others seem better suited for evaluation. For example, asking an SP for permission to perform an autopsy on the next of kin or informing an SP that a malignancy has recurred provides the student an opportunity to practice these very difficult communication skills in a safe setting. Common ambulatory problems are more appropriate for evaluating clinical skills such as data gathering. A very powerful use of this technique is to allow a student to interact with an SP and then to have the SP evaluate the student's performance and provide immediate feedback, discussing strengths and weaknesses in the student's performance. The SP can then return to his role, and the student can practice what he has just learned.

Many schools appear strongly interested in expanding their SP programs. A major barrier is the lack of trained staff to develop protocols and train SPs. The development of a network for sharing resources and conducting workshops on case development and training of SPs is now under way and should decrease this problem. We do not believe that SPs can ever replace the rich encounter that occurs between a student, a faculty member, and an actual patient, but they can be a valuable adjunct to teaching.

References

1. Stillman, P. L., Ruggill, J. S., and Sabers, D. L. The Use of Practical Instructors to Teach and Evaluate a Complete Physical Examination. *Eval. Health Prof.* **1**(1978): 49-54.
2. Stillman, P. L., Burpeau-DiGregorio, M. Y., Nicholson, G. I., Sabers, D. L., and Stillman, A. E. Six Years of Experience Using Patient Instructors to Teach Interviewing Skills. *J. Med. Educ.* **58**(1983)941-946.
3. Stillman, P. L., et al. Assessing Clinical Skills of Residents with Standardized Patients. *Ann. Intern. Med.* **105**(1986): 762-771.
4. Stillman, P. L., and Swanson, D. B. Insuring the Clinical Competence of Medical School Graduates Through Standardized Patients. *Arch. Intern. Med.* **147**(1987): 1049-1052.
5. Stillman, P. L., Regan, M. B., and Swanson, D. B. A Diagnostic Fourth-Year Performance Assessment. *Arch. Intern. Med.* **147**(1987):1981-1985.
6. Stillman, P. L., et al. An Assessment of the Clinical Skills of Fourth-Year Students at Four New England Medical Schools. *Acad. Med.* **65**(1990):320-326.
7. Kretzschmar, R. M. Evolution of the Gynecology Teaching Associate: An Education Specialist. Am. J. Obstet. Gynecol. **131** (1978):367-375.
8. Barrows, H. S. *Simulated (Standardized) Patients and Other Human Simulations.* Chapel Hill, North Carolina: Health Sciences Consortium, 1988.
9. Personal communication, Donald G. Kassebaum, M.D., director of accreditation, Association of American Medical Colleges, Washington, D.C., December 1989.

RESEARCH REPORTS

ARTICLES

The Problem of Late Hospitalization: A Quality and Cost Issue

JOSEPH S. GONNELLA, M.D., DANIEL Z. LOUIS, M.S., CARTER ZELEZNIK, Ph.D., and BARBARA J. TURNER, M.D.

Abstract—Emphasis on controlling health-care costs has led to many activities aimed at avoiding medically unnecessary hospitalizations. Much less attention has been given to patients hospitalized in advanced stages of illness and the impact of these late admissions on cost and quality of care. A panel of physicians developed criteria to categorize hospital admissions into one of three groups—early, timely, or late—based on the timing of the initial hospitalizations of patients admitted with any one of 14 diagnoses. Over a period of one year (fiscal year 1984) the criteria were applied retrospectively to 2,713 patients admitted to either of two hospitals. Twenty-one percent of the admissions studied in one hospital and 19% in the other were judged to occur later than was desirable. The mean length of stay for late hospitalizations exceeded that for timely hospitalizations by 11.1 days at one hospital and by 7.5 days at the other ($p < .01$). Similar patterns were observed in analyzing the 14 diseases individually and in an analysis of hospital charges at the one hospital where charge data were available. In-hospital mortality rates for patients with a principal diagnosis of bacterial pneumonia were over ten times greater for those admitted late than for those whose admissions were timely (39.0% versus 3.8%, $p < .001$, at one hospital; 28.9% versus 2.1%, $p < .001$, at the other). While not all late hospitalizations are avoidable, the authors believe that the analysis of late hospitalization patterns is an important part of any effort that can be made to reduce them. *Acad. Med.* 65(1990):314–319.

Current utilization review activities focus on medically unnecessary hospitalizations or those considered longer than necessary.[1-3] With this focus, these programs ignore or underestimate the problem of patients who are admitted to the hospital in advanced stages of diseases which potentially could have been detected and treated earlier. Changes in patient care reimbursement methods and in the structure of health care delivery organizations make it even more urgent to monitor hospitalizations of patients in late stages of disease.[4-6]

This study was designed to determine whether significant opportunities exist to reduce patient morbidity and health care costs by utilizing health care services before diseases progress to advanced stages.

Operationally, this study is based on the "disease staging" concept.[7,8] Staging is a disease classification system that focuses specifically on the biological dimension of severity of illness (etiology, pathophysiology, and ranking of local and systemic complications). Over the past ten years, staging criteria have been developed for 400 disease entities.[9] Disease staging previously has been used to evaluate the comparative costs of caring for patients with early and more advanced stages of disease. Analysis of over 30,000 Medicare discharges has shown that hospitalizations of advanced-stage diabetic patients cost 50% more than do those for patients in early stages.[10]

Method

In order to modify disease staging for the current project, a panel of physicians (selected from a number of departments at one medical school) divided the staging criteria for each of 14 diseases into three groups in terms of the timing of hospitalization:

1. *Early.* Accepted level of medical care can be provided in an outpatient setting. The admission is premature or inappropriate. (*Example*: admission of a patient with chronic cholecystitis who is treated only with oral medication and undergoes no surgical procedure.)
2. *Timely.* The patient is admitted at the earliest detectable stage at which the probability of disease progression can be significantly reduced through use of resources available only in a hospital setting; or the patient has another condition(s) unrelated to the principal diagnosis that makes ambulatory treatment inappropriate; or patient compliance is such that inpatient care is necessary. (*Example*: admission of a patient with acute cholecystitis without perforation of the gallbladder.)

Dr. Gonnella is senior vice president and dean, Jefferson Medical College, Philadelphia, Pennsylvania. Mr. Louis is managing director, Dr. Zeleznik is a senior researcher, and Dr. Turner is a research associate; all at the Center for Research in Medical Education and Health Care, Jefferson Medical College.

Correspondence and requests for reprints should be addressed to Mr. Louis, Jefferson Medical College, 1025 Walnut Street, Room 119, Philadelphia, PA 19107.

3. *Late.* Hospitalization is delayed. The patient has complications that could have been prevented with earlier hospitalization. (*Example*: admission of a patient with acute cholecystitis who has gangrene of the gallbladder.)

It is important to stress that not all late hospital admissions are preventable. Late hospitalization can result from patient-related problems (such as failure to seek care or failure to follow a prescribed treatment); provider problems (such as an incorrect or late diagnosis or treatment); or system problems (such as lack of access to care or an overzealous admission certification program). Also, in some cases, a disease may progress so rapidly that nothing could reasonably have been done to avoid the disease progression prior to hospitalization.

In this study the timing of hospitalization for 14 diseases was examined in two hospitals, a university-based hospital in a major metropolitan area (Hospital A) and a community-based hospital in a small city (Hospital B). Data were collected on all initial hospitalizations in fiscal year 1984 for 14 high-frequency diagnoses. These conditions were chosen to assure analysis of diseases involving a variety of body systems and a mix of medical or surgical treatments. Computer software was used with automated discharge abstract data to classify patients with any one of these 14 discharge diagnoses according to stage of disease and timeliness of hospital admission. Initial hospitalizations for the selected diagnoses totaled 1,549 for Hospital A and 1,164 for Hospital B.

A random sample of admissions to each hospital, stratified to assure representation of all 14 diagnoses and the three categories of timing of hospitalization, was selected for chart review to validate the diagnosis, staging, and hospitalization timing classifications assigned using the computer software. The sample included 532 admissions to Hospital A and 519 admissions to Hospital B.

The hospital medical records of these sample patients were abstracted using forms designed for each of the 14 sample diagnoses. Data were collected on patient demographic characteristics, admission and discharge diagnoses, preadmission medical history, laboratory evaluations, disease-specific complications, the presence of other diseases (comorbidities), and discharge status. The abstracted data were reviewed by physician consultants. Diagnosis, staging, and timing-of-hospitalization assignments were confirmed or, if necessary, changed on the basis of the physician review.

Using the sample data, a comparison was made between the combined mean lengths of stay for patients with one of the 14 selected diseases whose hospitalizations were early, timely, or late. Separate, within-hospital analyses were made between the mean lengths of stay for the timely and late categories. Comparisons were also made of hospital charges at Hospital A. (Analyses of hospital charges for patients in Hospital B were not performed because complete charge data were not available.) The *t*-test was performed to test the statistical significance of observed differences in the independent variables (timing-of-hospitalization categories) and in the means of the dependent variables (length of hospital stay and charges).

The full population data were used as the basis for an analysis of the rates of late hospital admissions for the 14 diagnoses and for an analysis of in-hospital mortality rates.

Results

Late hospitalizations for all patients with any one of the 14 diseases totaled 322 at Hospital A (21% of all patients treated for these diseases) and 231 at Hospital B (19% of all such patients).

The mean lengths of stay for timely and late hospitalizations in the stratified sample of the 14 diseases were compared for the two hospitals. For timely hospitalizations, the mean lengths of stay at Hospitals A and B were the same: 7.1 days. The mean lengths of stay for late hospitalizations were 18.2 days at Hospital A and 14.6 days at Hospital B.

The numbers of patients sampled from the timely and late groups, and the means and standard deviations of the lengths of stay at Hospital A, are presented in Table 1. For all 14 diseases the mean length of stay was longer for patients judged to be hospitalized late than for those whose hospitalizations were timely. Significant differences in lengths of stay were observed for five categories: bacterial pneumonia, cancer of the cervix, diabetes mellitus, diverticular disease, and urinary tract infection.

At Hospital A, five conditions, including external hernia, otitis media, pelvic inflammatory disease, purulent sinusitis, and uterine fibroma, had between 12 and 28 patients in the timely category but fewer than five patients in the late category. The small number of late hospitalizations makes statistical comparisons of mean lengths of stay between timely and late hospitalizations inappropriate. However, for these five diseases, the mean lengths of stay were also markedly longer for the late hospitalizations.

Combining all 14 conditions, the mean length of stay for late hospitalizations was 18.2 days, significantly exceeding the mean length of stay for timely hospitalizations, which was 7.1 days ($p < .01$).

The length-of-stay distributions of all timely and late hospitalizations at Hospital A for all 14 diseases combined are displayed in Figure 1. Less than 25% of the patients hospitalized on a timely basis had lengths of stay exceeding ten days, with very few patients (<10%) staying 20 days or longer. The length-of-stay pattern for the late hospitalizations was more highly skewed toward longer durations. Over half of these patients stayed more than ten days and over 20% stayed 20 days or longer.

Comparisons between mean lengths of hospital stay for the timely and late categories of the sample of patients who were admitted to Hospital B with any one of the 14 selected diseases are shown in Table 2. As in the analysis of mean lengths of stay for the patients at Hospital A, late hospitalizations were found to be consistently longer. Significant differ-

Table 1

Numbers of Patients, Mean Lengths of Stay (LOS), and *t*-Test Results for Timely and Late Hospitalizations in Hospital A, by Diagnostic Category, Fiscal Year 1984*

Diagnostic Category	Timeliness of Hospitalization						
	Timely			Late			
	n	Mean LOS	SD	*n*	Mean LOS	SD	*t*†
Appendicitis	21	9.6	15.4	16	12.9	8.5	n.s.**
Bacterial pneumonia	19	7.6	4.3	18	25.1	20.2	3.59¶
Cancer of cervix uteri	10	3.7	2.6	19	9.8	3.6	4.76¶
Cancer of colon and rectum	10	11.6	9.1	25	17.9	9.0	n.s.**
Cholecystitis	20	10.4	4.6	15	13.9	8.2	n.s.**
Diabetes mellitus	20	10.7	6.2	21	23.7	19.5	2.90¶‡
Diverticular disease	17	9.8	4.5	15	26.1	28.7	2.18‖§
External hernia	28	4.5	3.4	4	23.5	17.5	‡
Otitis media	19	3.5	3.2	2	8.5	7.8	‡
Pelvic inflammatory disease	12	4.6	1.5	0	—	—	‡
Peptic ulcer	22	6.9	5.3	20	11.2	9.2	n.s.**
Purulent sinusitis	25	5.3	5.3	2	53.5	.70	‡
Urinary tract infection	8	4.1	5.6	20	19.5	17.7	3.48¶§
Uterine fibroma	12	6.3	2.9	1	12.0	—	‡
All conditions	243	7.1	6.8	178	18.2	16.3	8.51¶§

*Over a one-year period, 1,549 patients admitted to Hospital A were classified according to whether their admissions were early, timely, or late, using criteria developed by a panel of physicians. This table compares data about a sample of the Hospital A patients whose admissions were timely with those of patients from the same hospital whose admissions were late, a total of 421 patients. The data show that the mean length of stay was longer for the patients who were hospitalized late, for all diseases. See Table 2 for comparable data from a sample of patients from Hospital B, the other hospital studied.

†The *t*-test for independent groups.

‡Insufficient sample size (fewer than five patients in one diagnostic group).

§Statistical test for homogeneity of variances was significant; therefore, *t*-test for unequal variances was applied.

‖$p < .05$

¶$p < .01$

**Not significant.

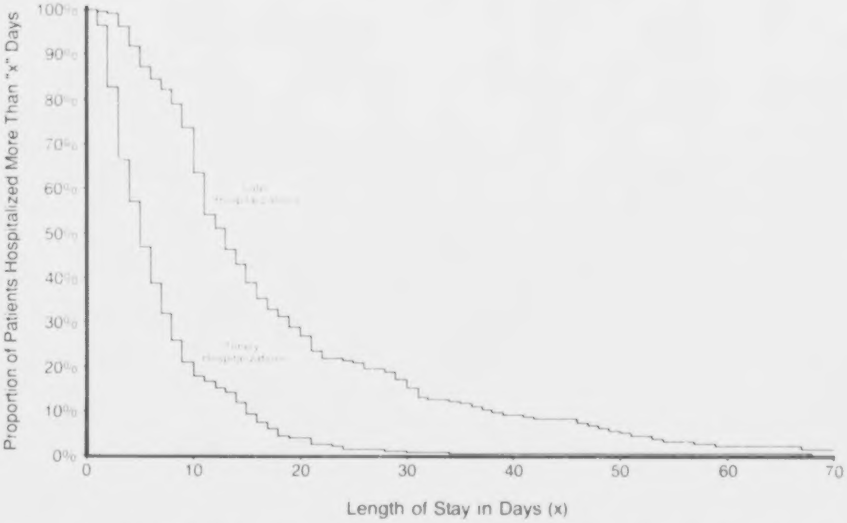


Figure 1. **The distribution of hospital lengths of stay for timely and late hospitalizations at Hospital A. The difference between the two curves indicates a higher probability of long stays for patients hospitalized late. Data from 532 admissions to Hospital A were used for this figure.**

ences between mean lengths of stay in the timely and late categories were observed for five of the diseases: appendicitis, bacterial pneumonia, cholecystitis, diverticular disease, and urinary tract infection.

As in the analysis of Hospital A data, small numbers of patients were found in the late-hospitalization group for six of the diseases, obviating statistical analysis. For each of these diseases—cancer of the cervix uteri, external hernia, otitis media, pelvic inflammatory disease, purulent sinusitis, and uterine fibroma—fewer than five patients had late hospitalizations. When mean length-of-stay data were determined for all 14 diseases combined at Hospital B, a significantly greater mean length of stay of 14.6 days was found for patients in the late-hospitalization category compared with 7.1 days for patients in the timely category ($p < .01$).

Table 2

Numbers of Patients, Mean Lengths of Stay (LOS), and *t*-Test Results for Timely and Late Hospitalizations in Hospital B, by Diagnostic Category, Fiscal Year 1984*

Diagnostic Category	Timeliness of Hospitalization						
	Timely			Late			
	n	Mean LOS	SD	*n*	Mean LOS	SD	*t*†
Appendicitis	24	3.2	1.0	16	8.9	3.3	6.89¶§
Bacterial pneumonia	19	8.2	4.9	20	17.0	14.8	2.70‖§
Cancer of cervix uteri	6	2.0	2.4	1	6.0	—	‡
Cancer of colon and rectum	17	19.9	12.3	23	20.0	12.6	n.s.**
Cholecystitis	22	8.6	2.8	19	13.6	8.0	2.59‖§
Diabetes mellitus	21	7.7	3.9	18	12.2	9.3	n.s.**
Diverticular disease	15	9.6	4.6	6	25.8	14.6	2.66‖§
External hernia	21	3.6	2.2	4	29.8	18.3	‡
Otitis media	11	4.3	3.0	2	5.5	.7	‡
Pelvic inflammatory disease	20	6.2	4.2	1	4.0	—	‡
Peptic ulcer	19	8.2	4.7	19	12.5	8.5	n.s.**
Purulent sinusitis	14	3.9	4.6	0	—	—	‡
Urinary tract infection	20	5.8	4.9	23	11.7	4.6	4.04¶
Uterine fibroma	18	6.5	2.5	0	—	—	‡
All conditions	247	7.1	6.2	152	14.6	10.9	7.70¶

*Over a one-year period, 1,164 patients admitted to Hospital B were classified according to whether their admissions were early, timely, or late, using criteria developed by a panel of physicians. This table compares data about a sample of Hospital B patients whose admissions were timely with those of patients from the same hospital whose admissions were late, a total of 399 patients. The data show that the mean length of stay was longer for the patients who were hospitalized late, for most diseases. See Table 1 for comparable data from a sample of patients from Hospital A, the other hospital studied.

†The *t*-test for independent groups.

‡Insufficient sample size (fewer than five patients in one diagnostic group).

§Statistical test for homogeneity of variances was significant; therefore *t*-test for unequal variances was applied.

‖$p<.05$.

¶$p<.01$.

**Not significant.

As displayed in Table 3, the mean charges for the Hospital A patients in the late category exceeded the mean charges for the same hospital's patients in the timely group for all diseases. For all conditions combined, the mean charge was $6,821 for patients hospitalized on a timely basis, compared with $20,728 for patients hospitalized late. (Hospital B data for these types of charges were not available.)

In eight of the diagnostic categories, and for all conditions combined, mean charges for late hospitalizations were significantly greater than mean charges for timely hospitalizations ($p < .05$). For three of these eight conditions, cancer of the colon and rectum, cholecystitis, and peptic ulcer, significant differences were observed in mean charges despite the fact that observed differences in mean lengths of stay were not significant for these diseases. Five of the diagnostic categories had too few late hospitalizations to justify a statistical test of significance.

An analysis of timely and late hospitalizations of two diseases, bacterial pneumonia and cholecystitis, offers examples of the wide range of hospital resource use within single diagnosis-related groups (DRGs). Eighteen patients with bacterial pneumonia in the sample population were grouped into DRG 89: "simple pneumonia and pleurisy, age greater than or equal to 70 and/or complicating conditions." When these patients were divided by timeliness of hospitalization, marked differences were evident in mean lengths of stay and mean charges. Eight patients whose hospitalizations were categorized as timely stayed 9.5 (SD 4.2) days, and their mean charge was $7,564 (SD $3,500); ten patients who were hospitalized late stayed 23.7 (SD 24.2) days, with a mean charge of $35,469 (SD $42,426).

Sixteen patients in the study population with cholecystitis were classified within DRG 197: "total cholecystectomy without common duct exploration, age greater than or equal to 70 and/or complicating conditions." Ten patients in the timely group stayed a mean of 9.8 (SD 3.9) days, with a mean total hospital charge of $9,549 (SD $3,326), in contrast to 13.3 (SD 11.4) hospital days with a mean charge of $15,602 (SD $10,646) for six patients in the late group.

Based on population data, the in-hospital mortality of all patients with a principal diagnosis of bacterial pneumonia shows an expected rise in mortality of the patients hospitalized in the late stage of the disease. At Hospital A, five of 132 patients

Table 3

Number of Patients, Mean Charges for Stay, and *t*-Test Results for Timely and Late Hospitalizations in Hospital A, by Diagnostic Category, Fiscal Year 1984*

Diagnostic Category	Timeliness of Hospitalization						
	Timely			Late			
	n	Mean Charges in $	SD	*n*	Mean Charges in $	SD	*t*†
Appendicitis	21	10,505	16,294	16	13,494	9,058	n.s.**
Bacterial pneumonia	19	6,546	5,398	18	38,533	35,087	3.82¶§
Cancer of cervix uteri	10	3,902	2,514	19	11,532	4,772	4.70¶§
Cancer of colon and rectum	10	10,341	7,700	25	18,652	8,843	2.60‖
Cholecystitis	20	10,415	4,849	15	15,649	7,993	2.25‖§
Diabetes mellitus	20	8,391	5,018	20	22,640	24,625	2.54‖§
Diverticular disease	17	8,971	4,967	15	31,901	38,135	2.31‖§
External hernia	28	4,202	2,208	4	23,097	15,554	‡
Otitis media	19	4,335	3,190	2	6,777	3,380	‡
Pelvic inflammatory disease	12	4,844	1,278	0	—	—	‡
Peptic ulcer	21	6,412	5,615	20	16,179	19,653	2.14‖§
Purulent sinusitis	25	5,284	4,613	1	70,621	—	‡
Urinary tract infection	7	3,795	3,690	20	18,687	19,014	3.33¶§
Uterine fibroma	11	6,312	2,803	1	13,081	—	‡
All conditions	240	6,821	6,705	176	20,728	22,195	8.05¶§

*This table compares the hospitalization charges of a sample of Hospital A patients whose admissions were timely with the charges of patients from the same hospital whose admissions were late, a total of 416 patients. The data show that the mean charges for the patients in the late category exceeded the mean charges for those in the timely group, for all diseases. Comparable data on charges for patients in these categories from Hospital B, the other hospital studied, were not available.

†The *t*-test for independent groups.

‡Insufficient sample size (fewer than five patients in one diagnostic group).

§Statistical test for homogeneity of variances was significant; therefore, *t*-test for unequal variances was applied.

‖$p < .05$.

¶$p < .01$.

**Not significant.

(3.8%) who were categorized as having timely hospitalizations died, whereas 16 of 41 patients (39.0%) in the late category died. Similar proportions were observed in the same categories for patients in Hospital B: three of 144 patients (2.1%) in the timely category died, whereas 11 of 38 (28.9%) in the late category died. (For both hospitals, $p < .001$.)

Discussion

With the current emphasis on cost containment, it is increasingly important to gain greater understanding of the factors affecting the cost of health care services. The principal aim of utilization review programs is to assure that patients using an acute care hospital are "sick enough" to justify the use of that expensive resource. The prospective payment system is designed to provide financial incentives for the efficient use of those resources through payment of fixed fees for treatment of predefined diagnostic types. The authors believe that there is a major gap in the focus of these programs. Neither traditional utilization review programs nor the prospective payment system adequately accounts for the patient who arrives at the hospital "sicker"—at a later stage—than is appropriate.

In this project, disease staging was applied to define objectively the severities of patients' diseases, in order to create categories of "isomorbid" patients. The case-mix classification system can also be differentiated in a way that will allow assessment of the appropriateness of the timing of a patient's hospitalization. Using the instruments developed for this project, data were collected for a sample of patients who had any one of 14 diagnostic categories from two hospitals to test the principal hypothesis: Significantly fewer resources are consumed by a patient hospitalized on a timely basis than are consumed by a patient hospitalized later than desirable in the disease process.

The data from this study's two sample hospitals strongly support this hypothesis. The differences between timely and late admissions in mean lengths of stay and mean charges (from the hospital for which charge data were available) were striking. For every diagnosis for which sufficient data were available to make meaningful comparison, the mean length of stay and mean charge for late hospitalizations exceeded the mean length of stay and mean charge for timely hospitalizations.

For all of the sample diagnoses combined, the mean length of stay for late hospitalizations exceeded the mean length of stay for timely hospitalizations by 11.1 days at one hospital and by 7.5 days at the other hospital. The mean charge per late hospitalization exceeded the mean charge per timely hospitalization by $13,907. For the 176 late hospitalizations included in the sample at hospital A, this differential amounted to over $2.4 million. This difference in charges resulted from the greater intensity of services (diagnostic and/or therapeutic) needed for these patients.

While the project did not focus on comparisons between the two sample hospitals, the proportions of late hospitalizations and the pattern of length of stay were quite similar in the two hospitals, indicating that the problem of late hospitalizations is not unique to large academic medical centers.

The intra-DRG analysis of bacterial pneumonia and cholecystitis patients highlights the heterogeneity found in a number of DRGs. Under a prospective payment system using DRGs for definition of case type, groups of patients who are consistently hospitalized late can be very costly to the hospital.

Insofar as late-stage treatment is costly and leads to increased risk of morbidity and mortality, it is important that factors related to delayed treatment be identified so that changes may be implemented. If a patient is treated only at a late stage of illness in a hospital, something has gone wrong. Possible explanations include

1. The disease has been treated appropriately on an outpatient basis but a lack of medical knowledge prevents its successful treatment.
2. The patient has not received appropriate medical treatment on an outpatient basis because of diagnostic or other physician errors.
3. The patient has not complied with treatment provided by the physician.
4. The patient did not seek early medical care for any of a variety of reasons, such as fear, ignorance, and lack of access to care.
5. The disease progression was so rapid or unobservable that the patient would have been unlikely or unable to seek early medical treatment.
6. The patient was hospitalized in a timely fashion but inpatient treatment was delayed or inappropriate.

Undoubtedly not all late hospitalizations observed in this study were preventable; however, hospitalizing even a small proportion of these patients at an earlier stage could have yielded substantial cost savings, as well as potential reductions in morbidity and mortality.

The techniques described in this paper can be used to identify patterns of inappropriately late hospitalization. Where such patterns are found, corrective action(s) focused on the identified problem areas can be implemented. Depending upon the contributions of the patient, the physician, the health care delivery system, and the community to these problems, appropriate activities might include specific patient education programs, patient screening programs, continuing medical education programs, and feedback to physicians and organizations that provide health care.[11]

Measurement of timing of hospitalization can also be used as a tool for the evaluation of the outcome of care provided in a variety of forms of outpatient programs, such as health maintenance organizations (HMOs). Theoretically, HMOs and other managed care organizations, through improved accessibility and continuity of care, should have proportionately fewer inappropriately early and late hospitalizations than would a fee-for-service system. Evaluating the patterns of timing of hospitalization of individuals enrolled in these types of plans in comparison with similar groups of patients would enable the testing of this type of hypothesis.

The dramatic changes in the organization and financing of the health care system necessitate the development and implementation of tools that can be used on a large scale to evaluate their impact on quality of care. The authors believe that the approach described in this paper can be used as a part of such evaluations.

The investigation reported in this article was supported by a grant from the W. K. Kellogg Foundation, Battle Creek, Michigan.

References

1. Siu, A. L., et al. Inappropriate Use of Hospitals in a Randomized Trial of Health Insurance Plans. *N. Engl. J. Med.* **315**(1986):1259-1266.
2. Sandrick, K. M. Blue Cross and Blue Shield of Michigan's Efforts to Change Practice Patterns. *ORB* **10**(1984):349-352.
3. Shahoda T. Preadmission Review Cuts Hospital Use. *Hospitals* **58**(August 1984):54-55.
4. Social Security Amendments of 1983, Public Law 98-21, 97 Stat., 65-172.
5. Schlesinger, M. On the Limits of Expanding Health Care Reform: Chronic Care in Prepaid Settings. *Milbank Mem. Fund Q.* **64**(1986):189-214.
6. *Report to Congress: Impact of the Medicare Hospital Prospective Payment System, 1984 Annual Report.* Department of Health and Human Services, Health Care Financing Administration, Office of Research and Demonstration, HCFA Pub. No. 03231. Washington, D.C.: Government Printing Office, 1984.
7. Gonnella, J. S., Hornbrook, M. C., and Louis, D. Z. Staging of Disease: A Case-Mix Measurement. *JAMA* **251**(1984): 637-644.
8. Gonnella, J. S., Louis, D. Z., and McCord, J. J. The Staging Concept: An Approach to the Assessment of Outcome of Ambulatory Care. *Med. Care* **14**(1976):13-21.
9. Louis, D. Z., Barnes, C. A., and Pepitone, T. *Disease Staging: A Clinically Based Approach to Measurement of Disease Severity.* Vols. 1-5, (PB83-254631). Rockville, Maryland: National Center for Health Services Research, 1983.
10. Conklin, J. E., Lieberman, J. V., Barnes, C. A., and Louis, D. Z. Disease Staging: Implications for Hospital Reimbursement and Management. *Health Care Financ. Rev. Supp.* [no volume no.] (November 1984):13-22.
11. Gonnella, J. S., Zeleznik, C. Factors Involved in Comprehensive Patient Care Evaluation. *Med. Care* **12**(1974):928-934.

Evaluation of Students in Medicine Clerkships

GREGORY J. MAGARIAN, M.D., and DENNIS J. MAZUR, M.D., Ph.D.

Abstract—This article reports the results of a survey of the 124 medical schools in the United States and Puerto Rico, mailed in late 1986; the majority of the responses from 101 schools were received in early 1987. The authors report their findings concerning the overall process used by these schools to evaluate medicine clerks, specifically identifying the importance given to subjective evaluations, made by attending physicians and/or housestaff, compared with the importance given to clerks' performances on objective means of evaluation based on written and/or oral examinations, whether national or faculty-developed. The study reveals that the subjective evaluations of clerks by faculty attending physicians and by housestaff working with the clerks were used by 98 and 83 of the responding schools, respectively, and were given the greatest evaluative weight. Examinations of all types were used by 87 of the responding schools. The National Board of Medical Examiners (NBME) Part II Medicine Examination was the most frequently used and heavily weighted examination, followed by written faculty examinations. Oral examinations were the least frequently used evaluative tools, carried the least weight, and were rarely used without another examination. In light of their findings, the authors maintain that further research should be done to clarify whether the NBME Part II Medicine Examination is an appropriate instrument to measure the knowledge acquired during clerkship experiences. *Acad. Med.* 65(1990):341–345.

Students enter their clinical years anxious about leaving the classroom environment, where the guidelines regarding the evaluation process have been well established. For the most part, success during the basic science years is measured by performance on examinations. Personal and professional attributes, ability to work well within a team, and the ability to work closely with patients and others are not necessarily significant factors in the evaluation of students during these years. However, these attributes are critical to success as a ward clerk, houseofficer, and physician. The ability to organize responsibilities to patients and other team members must be balanced against developing skill in patient write-ups for review and case presentations to attending physicians. Throughout their clinical years, students are subject to evaluations that are likely to be far more subjective than what they have previously experienced. Of great concern to students is that evaluations be fair and consistent and that they know what is expected for achieving different levels of performance.

This study was one component of a larger study based on a survey of medicine clerkship directors. Here we report the answers to four questions concerning the overall process of evaluating clerks as reported by medicine clerkship directors in the United States and Puerto Rico: (1) What role do subjective evaluations of attending physicians and/or housestaff play in the evaluation of medicine clerks? (2) What role does the performance of medicine clerks on written and oral examinations play in this evaluation process? (3) What types of examinations are used in this evaluative process? (4) What sources of guidance, if any, are given to medicine clerkship directors in determining how to evaluate medicine clerks?

Method

The authors surveyed 124 medical school medicine clerkship directors in the United States and Puerto Rico in the latter part of 1986; responses were received in early 1987. The questionnaire asked the directors to record the percentages of weight given to the following subjective and/or objective components used in grading medicine clerks: (1) subjective evaluations from faculty serving as attending physicians, (2) subjective evaluations from housestaff working directly with students, (3) the National Board of Medical Examiners Part II Medicine Examination (subsequently referred to as the NBME-med exam), (4) written examinations developed by faculty, and (5) oral examinations.

Results

Responses were received from 101 (85%) of the 124 medical schools surveyed. Table 1 lists the components and their relative weights used in grading reported by these schools, in descending order of the number of schools using each component. The five most frequently used components were those offered as choices on the questionnaire; the remainder were spontaneously reported by the schools. When a single percentage was given to estimate the weight of evaluations received from both attending physicians and housestaff, the authors gave each of these evaluation categories equal weight (that is, each was assigned half the total percentage that had been given the combined category). This assignment of equal weight was made for responses from 17 of the schools.

Table 2 lists the five major components from which grades were derived and categorizes each into percentiles of distribution. Lack of attending physician faculty input and lack of housestaff input into a student's final evaluation were reported by three and 19 schools, respectively. This table shows that attending physician faculty and housestaff evaluations carried the greatest weight in the evaluation of medicine clerks, followed

---

Dr. Magarian is director of the medicine clerkship and associate professor, and Dr. Mazur is associate director of the medicine clerkship and assistant professor; both are in the Department of Medicine, Oregon Health Sciences University, Portland. Both authors are also staff physicians in the Section of General Internal Medicine, the Portland Veterans Administration Medical Center, Portland, Oregon.

Correspondence and requests for reprints should be addressed to Dr. Magarian, VA Medical Center (111-P), 3710 SW U.S. Veterans Hospital Road, P.O. Box 1034, Portland, OR 97207.

Table 1

Averages and Ranges of Weights That 101 U.S. Medical Schools Assigned to 13 Components Used in Determining the Final Grades of Medicine Clerks, and the Number of Schools Using Each Component, 1986*

Component Used in Grading Clerks	Average of Weights Given by All Schools (%)†	Range of Weights (%)	No. Schools Using Component
Attending physicians' evaluations	42	10–100	98
Housestaff evaluations	33	10–85	82
National Board of Medical Examiners Part II Medicine Examination‡	27	10–65	57
Written faculty-developed examinations	25	5–50	40
Oral examinations	15	5–30	26
Chief of medicine evaluations	12	10–15	3
Objective Structured Clinical Examination	33	25–40	2
Tutors' evaluations	40		1
Practical clinical examinations	25		1
Professional habits and attitudes	25		1
Students evaluating each other	15		1
Written patient work-ups	15		1
Clinic performance	10		1

*The authors surveyed 124 medical schools in the United States and Puerto Rico in late 1986 concerning the importance given to (1) subjective and (2) objective (written and/or oral) means of evaluating medicine clerks. A total of 101 schools (80%) responded.

†When a single percentage was given to estimate the weight of evaluations received from both attending physicians and housestaff, the authors gave each category equal weight (that is, each was assigned half the total percentage that had been given the combined category). This assignment of equal weight was made for 17 of the responses.

‡One school used this component only to elevate the final grade when there was superior performance on the examination.

closely by the use of examinations. There was a direct correlation between the use of the five major components used in grading and the average weight given to that component of the grade. In other words, the frequency of use and the weight given to a specific evaluation component were both in descending order: attending physician faculty evaluations, housestaff evaluations, NBME-med exam performance, written examination performance, and oral examination performance. Other evaluative methods were reported very infrequently.

Table 3 reveals how frequently each of the five major components were used in various combinations in the evaluation of students. The most common means of evaluating students was the use of input from attending physician faculty and/or house officers who had worked with students during the clerkship (98 schools). Only 12 schools reported that their evaluations of student performance were exclusively from attending physician faculty and housestaff input, without use of any examination. In only one school was the entire evaluation and grade determined solely from faculty attending input. Seventeen (17%) of the 98 schools using faculty attending physicians' input in their evaluation processes did not use housestaff input in this process, whereas only one of the 82 schools using housestaff input in the process did not also use faculty evaluations.

A total of 87 of the 101 responding schools reported using one or more written examinations of some type. The NBME-med exam was used by 57 of these 87 schools (66%); faculty-developed written examinations were used by 40 schools (46%); and oral examinations administered by faculty were used by 26 schools (30%). A total of 55 of the 87 schools that reported using examinations used only one type of examination (63%). The NBME-med exam was used as the single examination by 29 of these 55 schools (53%); faculty-developed examinations were used in this way by 23 schools (42%); and oral examinations administered by faculty were used as the single examination by only three schools (5%). The Objective Structured Clinical Examination (OSCE) was used as the single examination in only one of the 87 schools.

The frequencies of the various combinations used to evaluate medicine clerks were reported as follows: overall, only one examination was used more frequently than were two or more (55 of the 87 schools using examinations, or 63%, versus 32 of these schools, or 37%). Of the 32 schools using two or more examinations, 15 (47%) used the NBME-med exam plus an oral examination; nine (28%) used the NBME-med exam plus a faculty-developed written examination; four (13%) used a faculty-developed written examination plus an oral examination; and four (13%) used all three of the just-mentioned examination types. In only four (13%) of the 32 schools using two or more examinations was the NBME-med exam not included.

Only two schools reported using the OSCE, although two others specifically noted plans to institute it during the following academic year, while yet others were considering it. Departmental chairpersons provided input that was assigned a specific percentage of a student's final grade in only three schools. It is interesting that in one school, students' evaluations of each other accounted for 15% of the medicine clerkship grade. Only one school reported deriving a specified percentage of students' grades from students' professional habits and attitudes (25%), while one other reported giving a specific weight (10%) to students' performances in outpatient clinics.

A single examination was given greater weight than cumulative faculty attending evaluations in 20 schools and greater weight than cumulative housestaff evaluations in 29 schools. Equal weight was given to examinations compared with cumulative faculty attending physician and housestaff evaluations of clerks in 13 and 14 schools, respectively. In both instances, the NBME-med exam was most commonly used, followed in descending order by faculty-developed written examinations and oral examinations. Overall, an examination carried greater or equal weight compared with cumulative evaluations from attending physician faculty or house officers in a third (34) of the responding schools. However, in two-thirds (67 of 101) of the responding schools, subjective assessments by faculty or housestaff working with clerks carried the greatest weight in determining final grades.

Review of how results from the NBME-med exam were used revealed no consistent means of using the information in assisting in the evaluation and grading of students. Fifty schools used the raw scores as the primary means of evaluating student examination performance. Typically, specific raw scores were established to separate different levels of performance, such as distinguishing honors from passing. The remainder of the schools used a variety of other means to interpret NBME-med exam performances by medicine clerks. In rank order, these included performance within the class or based on data from prior classes, performance based on the national percentiles of previous fourth-year students having taken the examination, a combination of raw score and performance within the class, a combination of national percentile and performance within the class, and (as reported by one school) a combination of raw score and national percentile.

Only 32 of the 101 responding clerkship directors reported that they had received any guidance in how to evaluate students, while 81 schools reported that faculty had received some form of instruction in the evaluation process. This was most commonly from memoranda and departmental meetings.

Table 2

Numbers of U.S. Medical Schools That Assigned Specific Weights to Five Major Components Used in Determining the Final Grades of Medicine Clerks, Listed by Percentile Distribution, 1986*

	Component Used in Grading Clerks				
	Evaluations of Attending Physician	Evaluations of Housestaff	NBME-med Exam†	Written Exams	Oral Exams
Percentile Distribution of Weights	No. Schools	No. Schools	No. Schools	No. Schools	No. Schools
0-10	2	5	2	8	12
11-20	4	13	15	10	6
21-30	22	22	22	10	6
31-40	30	23	9	7	—
41-50	21	12	4	4	—
51-60	6	—	—	—	—
61-70	1	2	—	—	—
71-80	1	1	—	—	—
81-90	1	1	—	—	—
91-100	1	—	—	—	—
Unknown	9	3	5	1	2
Total	98	82	57	40	26

*The authors surveyed 124 medical schools in the United States and Puerto Rico in late 1986 concerning the importance given to (1) subjective and (2) objective (written and/or oral) means of evaluating medicine clerks. A total of 101 schools (80%) responded.

†National Board of Medical Examiners Part II Medicine Examination.

Table 3

Numbers of 101 U.S. Medical Schools Using Specified Combinations of Components to Determine the Final Grades of Medicine Clerks, 1986*

Combination of Components Used in Grading Clerks†	No. Schools Using Combination
Attending physician, houseofficer, NBME-med exam	22
Attending physician, houseofficer, written exam	19
Attending physician, houseofficer, NBME-med exam, oral exam	12
Attending physician, houseofficer	11
Attending physician, houseofficer, NBME-med exam, written exam	7
Attending physician, NBME-med exam	7
Attending physician, houseofficer, written exam, oral exam	4
Attending physician, written exam	4
Attending physician, houseofficer, oral exam	3
Attending physician, NBME-med exam, written exam	2
Attending physician, NBME-med exam, written exam, oral exam	2
Attending physician, NBME-med exam, oral exam	2
All other combinations each used by only one school or indeterminable (one school)	6

*The authors surveyed 124 medical schools in the United States and Puerto Rico in late 1986 concerning the importance given to (1) subjective and (2) objective (written and/or oral) means of evaluating medicine clerks. A total of 101 schools (80%) responded.

†NBME-med exam = National Board of Medical Examiners Part II Medicine Examination.

Discussion

This study of the components used by departments of medicine and their clerkship directors in the evaluation of medicine clerks has revealed that the evaluation of medicine clerks is heavily based on subjective input provided by faculty and housestaff physicians. It is not known how variable or consistent the overall processes of evaluation by faculty and housestaff are. Faculty and housestaff will have worked for variable periods of time with medical students, typically on hospital inpatient services. Only one school reported the use of student performance in an outpatient clinic setting as a specific component of its overall evaluation of students.

The subjective evaluations of attending physicians and those of housestaff were the foundation of the evaluation process, being used by 98 and 82 schools, respectively. While 17 of 98 schools used faculty attending evaluations without evaluations from housestaff working directly with students, only one school reported using evaluations from housestaff without using evaluations by faculty in the evaluation of student performance on medicine clerkships. Furthermore, in only five schools of 81 reporting use of evaluations from both sources were the housestaff evaluations clearly weighted heavier than those from faculty, whereas the reverse was true in 22 schools. Therefore, overall, evaluations from attending physicians were the most frequently used and the most heavily weighted. These evaluations were also the most frequently used when only one evaluation component was used.

Because of the weight given to such input and the differences in backgrounds and the variable amounts of evaluative experiences of individual faculty members and housestaff, it is clear that all faculty and housestaff evaluating students should be well versed in departmental expectations and trained in proficiency in the evaluation process. In this regard, it is heartening that 81 medicine clerkship directors reported that their faculties had received some form of guidance in the student evaluation process. However, the authors of the present study believe memoranda or discussions in departmental meetings, the primary sources reported for such guidance, may be inadequate, and that workshops,[1] as reported by only 13 schools, may be far more effective, despite their labor intensity and potential expense. While one school reported the development of a new medicine clerkship curriculum to assure high-quality teaching, independent of teaching by housestaff and by attending physicians,[2] the authors do not believe it is feasible for most schools to develop an additional evaluative staff of housestaff and attending physicians to evaluate medicine clerks. This should be an endeavor shared by all attending physicians and housestaff working with clerks in inpatient or ambulatory care arenas.

Performance on examinations is also a major indicator of achievement used in the evaluation of medicine clerks. Others[3-5] have noted that clinical ratings alone are not adequate for measuring student progress on medicine clerkships. They argue that a comprehensive assessment of student performance in medicine clerkships should include written examinations as well as clinical ratings. Only 14 schools did not report using examinations of some form in their evaluations of clerks. For the three examinations used most often, the frequencies of their use reflected the relative weights applied to them.

The NBME-med exam was the examination most frequently used and most heavily weighted in the evaluation process according to responses from medicine clerkship directors in this survey. This may reflect less confidence in faculty-developed examinations and/or the need for more effort in their creation. However, despite its frequent use for evaluating medicine clerks, the NBME-med exam was not created for this purpose. Furthermore, to the authors' knowledge, no study has demonstrated its appropriateness in measuring clinical knowledge and judgment acquired during medicine clerkships. Performance on this examination at the end of medicine clerkships may reflect the overall accumulation of clinical knowledge and judgment up to that point in time in a student's total medical school experience, including other preceding clerkships. If such is the case, one might expect students to enter with a greater baseline medical knowledge base and perform better on the NBME-med exam when the medicine clerkship is taken later in the year, after they have had more clinical experiences from other preceding clerkships. If so, this needs to be taken into account to assure fairness in grading. The popularity of the NBME-med exam probably also relates to its reasonable cost ($25 per examination), ease of administration, development and standardization by professional test makers, and national normative values for comparing student performance.

Although over half of the responding schools used the NBME-med exam for clerkship testing, the normative values returned by the NBME did not include the one that would be most valuable, namely, how a school's medicine clerks performed in comparison with the medicine clerks from other schools taking the examination during the same period. Rather, the NBME-med exam provided normative data to compare a school's third-year clerks with a group of students who previously took the same examination when administered during their fourth year as part of the NBME examination — Part II. This is likely to have contributed greatly to the marked variability in the way individual schools used the results in their overall evaluation processes. For the NBME-med exam to continue to enjoy such popularity for use as an examination in medicine clerkships, its validity as an appropriate tool in this regard needs to be established. Although the weight given to performance on this examination is higher than that given to faculty-developed written or oral examinations, justification for this test's elevated stature has not been established. The authors are unaware of data on the correlation between performances of medicine clerks on the NBME-med exam and concomitant performances on written or oral examinations. A previous study[4] revealed poor correlation between NBME-med exam performance and the evaluations of faculty who had worked with the same students, and suggested that a halo effect may be playing a role. This latter effect would not be expected to affect written examinations or oral examinations, if conducted by examiners unaware of a student's performance as perceived by others.

Oral examinations were used as the only examination by only three schools (12% of schools using single examinations), which probably reflects their labor intensity and an overall distrust of their use as a consistent and fair measure of students' knowledge and judgment. In only four schools were oral examinations paired with another faculty-developed written examination. In other words, when oral examinations were used, 70% of the time they were used in conjunction with the standardized NBME-med exam.

Summary

The results of this survey strongly suggest that faculty physicians and housestaff are the most heavily relied upon to evaluate medicine clerks. It would seem important to assure that they are adequately trained to do so beyond simply having received guidelines from memoranda and discussions in departmental meetings. Workshops may be particularly effective in this regard.

Likewise, if the NBME-med exam is to be used in clerkship evaluations, it is important to establish whether it measures what is wanted. When taken by most students in their fourth year, it provides a reflection of where a student is, relative to his peers, locally and nationally, at that point in time, based on multiple clinical experiences beyond the medicine clerkship. The question remains: does performance on the NBME-med exam provide a useful indicator of knowledge acquired during the clerkship, or does it primarily provide a measure of a student's knowledge based on all his or her medical school experiences from both basic science and clinical studies? If the latter is true, use of the NBME-med exam may not be appropriate as a tool for evaluating the development of students during medicine clerkships. Similarly, if faculty-developed written or oral examinations are used, training in their development and use would seem equally important, to assure that they also measure what is desired. The authors believe a major factor leading to the relative lack of use of faculty-developed examinations is that faculty are uncertain of their skills in the development of such examinations and of whether such examinations are fair and consistent. This uncertainty may be a reflection of the paucity of clerkship directors who have received specific training in the evaluation of clerks. The authors believe that for improvement to occur in the overall evaluation process of students on medicine clerkships, it is important that faculty and housestaff receive more and better training in the evaluation process. However, since the minority of medicine clerkship directors reported having such training themselves, it is not clear how and by whom this training would occur and who (medical schools versus individual departments) would fund it.

References

1. Vadiserri, R. O., Koziol, S. M., Korth, W. W., and Haynes, L. T. A Workshop to Improve the Teaching Skills of Physician Faculty Members. *J. Med. Educ.* **61**(1986):469-471.
2. Regan-Smith, M. G. Improved Teaching Within a Medicine Clerkship. *J. Med. Educ.* **59**(1984):753-755.
3. Ramsey, P. G., et al . Use of Objective Examinations in Medicine Clerkships: Ten-Year Experience. *Am J. Med.* **81**(1986):669-674.
4. Marienfeld, R. D., and Reid, J. C. Subjective vs. Objective Evaluation of Clinical Clerks. *N. Engl. J. Med.* **302**(1980):1036-1037.
5. O'Donohue, W. J., Jr., and Wergin, J. F. Evaluation of Medical Students During a Clinical Clerkship in Internal Medicine. *J. Med. Educ.* **53**(1978):55-58.

(WONCA).

This new model has had a positive impact in Malta. Through WONCA, the model has attracted the attention of China and several countries in Latin America. In China it was discussed at the Chinese Medical Association's First International Scientific Meeting of General Practice in November 1989.

Reference

1. College of Family Physicians of Canada. *Educational Objectives for Certification in Family Medicine.* Toronto, Canada: CFPC, 1981.

SHORT PAPER

The Availability of Fellowship Training for Foreign Family Physicians

MICHAEL D. COHEN, M.B., Ch.B., and JOEL H. MERENSTEIN, M.D.

Abstract—The authors surveyed the 49 U.S. fellowship programs in family medicine listed in the 1988 *Fellowship Directory* published by the Society of Teachers of Family Medicine. The past experience and future intentions of the 47 responding programs toward non-immigrant foreign family physicians coming to the United States to undergo fellowship training in academic family practice were studied. Only a minority of the programs (22%) had ever trained a foreign fellow, and of these only five of 11 were definitely able to do so in the future. Over half (52%) of the programs that had never trained a foreign fellow stated that they were not prepared to train a foreign physician. Although foreign family physicians are encouraged by the American Academy of Family Practice to come to the United States for fellowships, this study shows that it is extremely difficult for these physicians to find programs willing to act as host. *Acad. Med.* 65(1990):348–350.

Although fellowships in academic family practice exist in the United States, they are uncommon in other countries. Letters written to the representative family practice organizations in seven English-speaking countries of the world (Australia, Britain, Canada, Ireland, New Zealand, South Africa, and the United States) revealed that there is no directory of institutionally funded academic family practice fellowships outside the United States and that few countries have any organized program to prepare family physicians for a career in academic family practice.

Apparently aware of this, the American Academy of Family Physicians has established a policy to encourage foreign physicians to come to the United States to study family practice and return to promote it in their home countries.[1]

The purpose of this study was to evaluate the experience and availability of fellowship training for foreign family physicians in the United States.

Method

The authors sent a letter to each of the fellowship programs listed in the 1988 *Fellowship Directory* published by the Society of Teachers of Family Medicine[2] asking whether they had trained any foreign fellows in their departments. A stamped, self-addressed postcard was provided for the reply. A foreign fellow was defined as a physician who fulfilled the following criteria: (1) was born outside the United States and is not a U.S. citizen; (2) has not studied at a medical school in the United States; (3) has undergone residency training in family practice outside the United States; (4) has undergone a fellowship program in academic family practice in the United States of at least one

---

Dr. Cohen is director of Be'ere Health Center of Kupat Cholim, Netanya, Israel. At the time of this study, he was a fellow in family practice at St. Margaret Memorial Hospital, Pittsburgh, Pennsylvania. Dr. Merenstein is director of research and fellowships in the family practice residency, also at St. Margaret Memorial Hospital.

Correspondence and requests for reprints should be addressed to Dr. Cohen, Director, Be'ere Clinic, Be'ere Street, Netanya, Israel.

A related article, "A New Faculty Training Model for Countries Developing Academic Family Medicine Programs," appears on page 346.

year's duration; and (5) intends to return to his or her home country at the end of the fellowship and has not come to the United States as an immigrant.

When a reply was received, a second, more detailed questionnaire was sent. If the program had trained a foreign fellow in the past then the questionnaire asked whether the organizers intended to train foreign fellows in the future. A second questionnaire was also included concerning each of the fellows who had been or were being trained, asking specific questions concerning the fellow's country of origin and other aspects of the person's history to check whether the criteria for a foreign fellow, listed in the previous paragraph, had been met. If the program had not trained a foreign fellow in the past, then the questionnaire asked whether it was possible for foreign physicians to pursue fellowships at that institution and, if not, what the reasons for excluding them were. If a program did not reply at either stage of the survey, then, after a month had elapsed, the appropriate questionnaire was sent again. If the second questionnaire mailing did not evoke a response, then the program director was telephoned personally and the questions asked directly.

Many of the fellowship programs listed in the *Fellowship Directory* were no longer in existence and the correspondents and telephone numbers for some others were often quite different from the ones given in the directory.

Results

Initial responses were received from 47 of the 49 programs listed. Of these, 36 (77%) had never had a foreign fellow participate in their programs. For the second questionnaire a total of 46 programs responded: ten of the 11 programs that had had foreign fellows, and all of the 36 who had not trained foreign fellows.

These data show that only five of the ten programs that had trained foreign fellows in the past were sure that they would be able to do so in the future. Of the 36 programs that did not have a foreign fellow, ten would be prepared to take foreign fellows in the future. The largest number of programs had never trained a foreign fellow and did not intend to do so in the future. Of the 46 programs responding, 21 said they might consider a foreign fellow; of these 21, six were uncertain of having funding available for foreign fellows.

Table 1

Origins of the 22 Former or Current Foreign Family Practice Fellows in the United States, 1988*

Origin	No. of Fellows	% of Fellows
Country		
Israel	5	23
Chile	3	14
Canada	2	9
United Kingdom	2	9
South Korea	2	9
Venezuela	2	9
Costa Rica	1	4.5
Japan	1	4.5
Columbia	1	4.5
Taiwan	1	4.5
New Zealand	1	4.5
Laos	1	4.5
Total	22	100
Area of world		
Latin America	7	31.8
Middle East	5	22.7
Far East	5	22.7
British Commonwealth†	3	13.7
Canada	2	9.1
Total	22	100

*In 1988 the authors surveyed the 49 U.S. fellowship programs in family medicine to investigate the past experience and future intentions of such programs toward non-immigrant foreign family physicians coming to the United States for fellowship training, and to obtain information about the origins of the fellows who had been, or were being, trained at that time. A total of 47 of the programs responded.

†British Commonwealth countries, except for Canada, are classified together.

The reasons programs gave for not considering foreign applicants were: (1) that a medical license would be unobtainable (three programs); (2) that there could be potential linguistic and communications problems (one); (3) that the program no longer existed (three); (4) that American candidates were always better (six); and (5) that the program's funding specifically precluded non-U.S. citizens (five).

Information was also obtained about the countries of origin of the fellows. Table 1 shows this information by country of origin and region of the world (although British Commonwealth countries other than Canada are classified together). The 22 fellows identified in this study were from all over the world apart from the African and Eastern-bloc countries.

Discussion

Although 21 of the 46 fellowship programs that responded to the second questionnaire stated that they would be willing to accept foreign trainees, six of those 21 were uncertain of their funding for foreign fellows. Only six had prior experience and a willingness to accept future trainees.

It is, however, encouraging that almost all of the programs with prior experience remained positive about accepting future fellows from other countries if funds were available.

No information was available on how many foreign family physicians apply for fellowships in the United States or what the perceived need for this training is in various countries.

Training teachers of family medicine has been a major goal of family practice organizations in this country, with the intention of improving the quality and quantity of practicing family physicians. This need is as great in other countries, but their resources for training are much more limited.

This study identified two major understandable explanations for not accepting foreign fellows: a program having funding that specifically precludes non-U.S. citizens and a lack of adequate secure funds. Other reasons are less understandable. Language and licensing concerns may be referred to the Educational Commission for Foreign Medical Graduates (ECFMG), which has been specifically charged with the task of dealing with them. That U.S. graduates are always of higher quality is clearly an unjustified bias and misses the point of accepting foreign fellows. Also, the unique contribution the visiting fellow brings to the host program in representing different ideas and perspectives can be considerable.

The International Medical Scholars Program, co-ordinated by the ECFMG and sponsored by many of the most prestigious U.S. medical organizations, including the American Medical Association and the American Hospital Association, has recently been established. Its functions are expected to be those of reviewing and approving applications, facilitating placements, and providing stipend support where needed, after fellowship funds become available. This program should expedite the training of foreign family practice faculty but will be able to do so only if there are programs in the United States that are willing to accept these physicians. Our data indicate that at present there is a serious shortage of such programs.

References

1. American Academy of Family Physicians. *1985 AAFP Transactions.* Kansas City, Missouri: AAFP, 1985, p. 188.
2. Society of Teachers of Family Medicine. *Fellowship Directory.* Kansas City, Missouri: STFM, 1988.

13273-8

JUNE 1990

ACADEMIC MEDICINE

Class Mentors at Wisconsin

Innovations in Clinical Practice Through Hospital-Funded Grants

Machiavelli and the Chief Resident

Academic Medicine and the NBME

Book Reviews

Geriatrics Education

In Residency Programs

In Nursing Homes and CCRCs

With Multidisciplinary Teams

Assessing Clinical Skills of FMGs

Gender and the Medical School Interview

VOLUME 65 • NUMBER SIX

JUNE 1990
VOLUME 65
NUMBER SIX

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Invited Articles

Class Mentors: A Step toward Implementing the GPEP Report, *by Charles C. Lobeck, M.D., and Howard L. Stone, Ph.D.* 351

Innovations in Clinical Practice through Hospital-funded Grants, *by Patricia D. Franklin, M.D., M.B.A., Robert J. Panzer, M.D., Leo P. Brideau, M.H.A., and Paul F. Griner, M.D.* 355

Machiavelli and the Chief Resident, *by Mario C. Raviglione, M.D.* 361

Ideas for Medical Education

The Problem-based Curriculum at Bowman Gray School of Medicine, *by James R. Philp, M.D., and Martha G. Camp, Ph.D.* 363

Book Reviews

A Year-Long Night, 364; *Doctors: The Biography of Medicine,* 365; *Flying Without Wings: Personal Reflections on Being Disabled,* 366; *The Rights of Patients: The Basic ACLU Guide to Patient Rights,* 367; *Assessing the Impact of Continuing Medical Education through Structured Physician Dialogue,* 368; *The Roots of Modern Biochemistry: Fritz Lipmann's Squiggle and Its Consequences,* 369; *Thomas Jefferson University: Tradition and Heritage,* 370; *Medical Revolution in Minnesota: A History of the University of Minnesota Medical School,* 371; *Disease and Discovery: A History of the Johns Hopkins School of Hygiene and Public Health, 1916–1939,* 372; *History of Medicine,* 373; *The History of Surgery in the United States, 1775–1900, Volume I: Textbooks, Monographs, and Treatises,* 374; *Maimonides' Commentary on the Aphorisms of Hippocrates,* 375

From the President 376

AAMC Focus 377

National Policy Perspectives 378

Editorial

Academic Medicine and the National Board: A Love-Hate Relationship? *by Edward J. Stemmler, M.D.* 380

Research Reports

ARTICLES

Geriatrics in Residency Programs, *by David B. Reuben,* M.D., *Arlene Fink,* Ph.D., *Susan Vivell,* Ph.D., *Susan H. Hirsch,* M.P.H., *and John C. Beck,* M.D. 382

Postbaccalaureate Preparation and Performance in Medical School, *by Mohammadreza Hojat,* Ph.D., *Robert S. Blacklow,* M.D., *Mary Robeson,* M.S., *J. Jon Veloski,* M.S., *and Bette D. Borenstein,* Ph.D. 388

Evaluating the Clinical Skills of Foreign Medical School Graduates Participating in an Internship Preparation Program, *by Arthur I. Rothman,* Ed.D., *Robert Cohen,* Ph.D., *and John Ross,* M.D. 391

Assessing Learning-Style Inventories and How Well They Predict Academic Performance, *by Lisa I. Leiden,* Ph.D., *Ross D. Crosby,* Ph.D., *and Hugh Follmer,* M.D. 395

The Nature of Patients' Requests for Physicians' Help, *by Sherman Eisenthal,* Ph.D., *Cheryl Koopman,* Ph.D., *and John D. Stoeckle,* M.D. 401

Faculty Development and Organizational Systems Behavior, *by Charles E. Henley,* D.O., M.P.H., *and David Magelssen,* M.D., M.A. 406

SHORT PAPERS

The Influence of Applicants' Gender on Medical School Interviews, *by Jill A. Marquart, Kathleen N. Franco,* M.D., *and Brendan T. Carroll,* M.D. 410

The Nursing Home As a Site for Teaching Medical Students, *by Myra Wiener,* M.D., *and Ann Shamaskin,* M.D. 412

The Emerging Relationship between Academic Health Centers and Continuing Care Retirement Communities, *by H. Garland Hershey,* D.D.S., *Mark E. Williams,* M.D., *and James C. Sadler,* Ph.D. 414

Teaching the Multidisciplinary Team Approach in a Geriatrics Miniresidency, *by J. Edward Jackson,* M.D., *Wigbert Wiederholt,* M.D., *and Robert Katzman,* M.D. 417

Brief Communications 419

AAMC Data Report Results of the NRMP for 1990 424

Publishing Services: **Hanley & Belfus, Inc.**
210 South 13th Street, Philadelphia, PA 19107

JUNE 1990
VOLUME 65
NUMBER SIX

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
Edward J. Stemmler, M.D.
Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
Arizona Health Sciences Center

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
Georgetown University Hospital
Philip Anderson, M.D.
University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
Harvard Medical School
Eta Berner, Ed.D.
University of Alabama School of Medicine
Ruth Bulger, Ph.D.
National Academy of Sciences
Vincent A. Fulginiti, M.D.
Tulane University School of Medicine
Nancy Gary, M.D.
Albany Medical College
Beatrix Hamburg, M.D.
Mt. Sinai Hospital
Murray M. Kappelman, M.D.
The University of Maryland School of Medicine
Ralph W. Muller
University of Chicago Hospitals
Herbert Pardes, M.D.
College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
University of Alabama School of Medicine
W. Vickery Stoughton
Toronto General Hospital
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
David H. Cohen, Ph.D.
Northwestern University

Chairman-Elect
William T. Butler, M.D.
Baylor College of Medicine

Immediate Past Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center School of Medicine

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Joe Dan Coulter, Ph.D.
University of Iowa College of Medicine

Chairman, Council of Teaching Hospitals
Raymond G. Schultze, M.D.
University of California, Los Angeles Medical Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.

Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036; (202) 828-0416.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

RESEARCH REPORTS

ARTICLES

Geriatrics in Residency Programs

DAVID B. REUBEN, M.D., ARLENE FINK, Ph.D., SUSAN VIVELL, Ph.D.,
SUSAN H. HIRSCH, M.P.H., and JOHN C. BECK, M.D.

Abstract—To determine how well geriatrics has been integrated into residency training, the authors surveyed a random 33% sample of all the 378 family practice ($n = 126$) and 420 internal medicine ($n = 140$) training programs in the United States in 1988. All the programs responded. On average, the internal medicine programs had more geriatrics faculty than did the family practice faculty, but these numbers were insufficient to meet current or future needs. Fewer than half of the residencies had geriatrics inpatient or ambulatory-care evaluation units, geriatrics consult services, geropsychiatry wards, or geriatrics clinics available as training sites. In contrast, nursing homes were available for 93% of the family practice programs and 58% of the internal medicine programs. A total of 80% of the family practice programs but only 36% of the internal medicine programs had geriatrics curricula in place. The authors conclude that integration of geriatrics content into residency training is far from universal, largely because of a shortage of faculty and clinical training sites. *Acad. Med.* 65(1990):382–387.

Increased life expectancy and decreased birth rates have led to dramatic absolute and relative increases in the numbers of persons who are 65 years of age and older. In 1988, 12% of the population (30.4 million) were 65 or older; by 2030, the percentage will increase to nearly 20% (65.6 million).[1] This growth has important implications for the health care system because older Americans consume health care resources in excess of their representation in the population. Currently, the 12% of persons over 65 account for 20% of physicians' office visits, 25% of medications, and 40% of acute hospital admissions.[2]

Medical educators have responded to the aging of the American population by developing specific programs to teach physicians to care for the elderly. In the eight-year period from 1976 to 1984, the number of U.S. medical schools providing undergraduate clinical training rose from 15 to 88.[3,4] The number of geriatric medicine fellowship programs almost tripled, from 36 to 1980 to 103 in 1987.[5] Geriatrics-related continuing medical education programs have increased from an average of 14 per year in the mid-1970s to 43 per year in the mid-1980s.[5]

Of all the stages of medical education, residency training is perhaps most pivotal in preparing physicians to practice medicine. It is a time of intense clinical experience and, for many, the last formal training before practice. For others, it is the last general training they will receive before concentrating on subspecialty medicine.

Training in geriatric medicine during residency emerges as a particularly important issue because the majority of clinical care has been and will continue to be delivered by physicians without specific training in geriatrics. In 1985, over half of all ambulatory care of persons 75 years of age and older was provided by general or family practitioners and internists.[2] An analysis of 1986 Medicare data reveals that physicians in these two specialties provide almost two-thirds of hospital care and over 95% of nursing home care of elderly persons.[6]

In response to the changing demographic profile and the role of primary care physicians in providing clinical care to elderly persons, special requirements have been issued in both internal medicine and family practice to increase residency training in geriatric medicine.[7–9] The resulting growth of specific geriatrics training in internal medicine (IM) and family practice (FP) residency programs has been reported in several studies. In 1979, Robbins and colleagues[10] reported that 21% of the 28 academic geriatrics units, nationwide, offered mandatory housestaff rotations. By 1984, Calkins[11] noted that 43% of 30 academic units provided mandatory housestaff rotations for most of the residents in internal medicine and/or family practice, almost a doubling of residency training opportunities in geriatrics. A 1980 survey[12] of family medicine residency programs reported that less than one-

---

Dr. Reuben is assistant professor, Dr. Beck is professor, and Dr. Vivell and Ms. Hirsch are staff at the Multicampus Division of Geriatric Medicine and Gerontology, University of California at Los Angeles School of Medicine. Dr. Fink is adjunct professor of medicine and public health, Department of Medicine, UCLA School of Medicine and UCLA School of Public Health.

Correspondence and reprint requests should be addressed to Dr. Reuben, Multicampus Division of Geriatric Medicine and Gerontology, UCLA School of Medicine 32-144 CHS 10833 Le Conte Avenue Los Angeles, CA 90024-1686.

Other articles about geriatrics training appear on pages 412, 414, and 417.

fourth had a geriatrics teaching program in place and only 29% had plans for developing such a rotation. By 1985, 76% of family practice residency programs offered specific training in geriatrics, and in 81% of these, the training was mandatory.[13] An unpublished 1985 survey[14] by the Association of Program Directors in Internal Medicine found that 60% of programs in internal medicine offered formal teaching in geriatrics, though this teaching may have amounted to only an occasional lecture. According to that survey, internal medicine residents were nine times more likely to take an elective rotation in cardiology than in geriatrics. Still lacking, however, is a comprehensive overview of geriatric medicine training during internal medicine and family practice residencies that evaluates and compares the two specialties that provide the majority of care to older persons.

To determine how well geriatrics has been integrated into residency training, we conducted a national survey of residency training programs in internal medicine and family practice.

Method

The survey was designed to gain a national perspective on current and planned educational activities in geriatrics at the residency training level.

Sampling

Using the 1988–1989 *Directory of Graduate Medical Education Programs*,[8] we randomly chose for the survey a 33% sample of all the 420 IM and 378 FP programs (excluding military programs) in the United States. Because of wide variation in the structures of FP programs, before we made our random choices of programs, we stratified the programs on the basis of type of program (medical school-based; community-based, medical school-administered; community-based; medical school-affiliated, community-based). The survey sample consisted of 126 FP programs and 140 IM programs. Of the IM programs, 29 (21%) had primary care tracks.

Survey Description

We developed a survey instrument (available on request from the authors) that asked about program characteristics; need and supply of geriatrics faculty (defined as faculty with expertise or training in geriatric medicine); training sites; the value of specific training in geriatric medicine; obstacles to the development of geriatrics curricula; and characteristics of the specific geriatrics curriculum, if present or planned. We defined "current and planned geriatrics curriculum" broadly to include clinical and didactic teaching; a subsequent question asked what percentage of the curriculum was clinical. We also asked whether the institution had specific programs for faculty development in geriatrics.

The survey questionnaire was structured to distinguish three levels of program development: geriatrics curriculum already in place, geriatrics curriculum planned within five years, and geriatrics curriculum not in place nor planned within five years. All who responded completed a common part of the survey and then, depending on the state of development of their programs, they were directed to complete appropriate remaining parts of it. Opinion questions used a three-point Likert scale (very important, important, not important) to rate the importance of specific elements of geriatrics training. The survey was tested by faculty members of the UCLA Multicampus Division of Geriatric Medicine and revised before being administered nationally.

Survey Administration

Prior to the administration of the survey, the project was given the support of the Association of Program Directors in Internal Medicine (APDIM) and the Association of Departments of Family Medicine (ADFM). This support was conveyed in a cover letter mailed with the survey to the program directors in March 1988. The survey was completed by the program directors or forwarded to an appropriate faculty member, often a geriatrician. Although confidentiality was assured, the name and title of the person completing the questionnaire were recorded to allow comparisons between geriatricians (defined by a title that indicated an administrative role in the geriatrics curriculum or a name known by the investigators to be that of someone who had completed a geriatrics fellowship or its equivalent) and non-geriatricians.

We gave the physicians one month to return the completed survey by mail and then called those who had not responded and completed their surveys by telephone. The telephone interviewers were trained by the research team and monitored closely to insure that the information collected by telephone was comparable to that collected by the mailed-in surveys. In addition, the responses to key questions were compared on the basis of whether the survey had been received by mail or completed over the telephone.[15] In general, there were no significant differences between the responses received by mail and those taken over the telephone.

Statistical analysis included descriptive statistics, chi-square, analysis of variance (ANOVA), and t-tests using the Bonferroni modification to adjust for multiple testing.

Results

All the 266 sampled programs responded. One hundred and seventy-one (66%) of the responses were received by mail and the remainder were taken by telephone. Over half (53%) of the respondents were departmental chairs or program directors, while 47 (18%) were geriatrician faculty members; nine (3%) were curriculum directors for residency program; 41 (19%) were other staff members (for example, administrative coordinators) and 16 (6%) were unidentified.

The vast majority of programs (87% in internal medicine and 90% in family practice) were affiliated with medical schools. The family practice programs were smaller (on average, 19 total residents) than were the categorical IM programs (on average, 41 total residents) and were approximately the same size as the primary

Table 1

Average Numbers of Geriatrics Faculty Available and Needed in 126 Family Practice Programs and 140 Internal Medicine Programs, 1988*

Type of Faculty	Average Number of Faculty: Internal Medicine	Average Number of Faculty: Family Practice
Available		
Full-time geriatrics faculty	2.8	2.3
Full-time geriatrics faculty with formal training in geriatrics	0.6	0.4
Part-time or voluntary geriatrics faculty	4.4	2.8
Needed		
Full-time geriatrics faculty currently needed	4.0	2.2†
Full-time geriatrics faculty needed in five years	5.2	2.9†

*To determine how well geriatrics has been integrated into residency training, the authors surveyed and obtained responses from 33% of all 378 family practice programs and all 420 internal medicine programs in the United States.

†$p < .05$ between specialties.

care IM programs (on average, 18 total residents).

Geriatrics faculty resources are presented in Table 1. Although the IM programs had more geriatrics faculty, the ratio of residents per full-time geriatrics faculty was higher in internal medicine (13.2:1) compared with family practice (7.4:1). Furthermore, 33% of the IM programs and 13% of the FP programs had no full-time faculty with expertise or training in geriatrics ($p < .001$). A minority of the programs in both disciplines had formally trained geriatricians as full-time faculty members; 65% of the IM programs and 71% of the FP programs had no formally trained full-time faculty member in geriatrics (not significant). Fewer than one-third of the programs in either specialty offered specific programs in geriatrics faculty development.

Existing and projected clinical sites for teaching geriatrics are presented in Table 2. Fewer than half of the IM and FP residencies had inpatient or ambulatory geriatrics evaluation units (GEUs), inpatient geriatrics consult services, geropsychiatry wards, or geriatrics clinics currently available as training sites. Even if all anticipated development of training sites is achieved, the majority of programs in both specialties will not have geropsychiatry wards available within the next five years. Significant differences between specialties were evident in current availability of inpatient GEUs (more common in the IM programs) and nursing homes (more common in the FP programs).

Fifty-one (36%) of the IM programs and 101 (80%) of the FP programs had geriatrics curricula in place at the residency training level ($p < .001$). In internal medicine 61% of curricula were required and in family practice 92% were required. Current curricula in both specialties were, on average, approximately five weeks long, and about three-fourths of the time was spent in clinical experiences. An additional 42% of the IM programs and 15% of the FP programs planned to implement geriatrics curricula within five years (difference not significant). Almost all those who responded envisioned the scopes of their programs in geriatrics to be greater in the next five years.

We compared the programs that had geriatrics curricula in place with those that did not with respect to faculty resources and training sites (Tables 3 and 4). Faculty resources and needs were similar within disciplines, except that the IM programs with geriatrics curricula had more formally trained geriatricians. Marked differences between the IM and FP programs with geriatrics curricula were apparent, however. Of note, 42% of the IM programs and 68% of the FP programs with curric-

Table 2

Numbers and Percentages of 126 Family Medicine Programs and 140 Internal Medicine Programs with Clinical Sites for Teaching Geriatrics and with Plans for Such Sites, 1988*

Type of Clinical Site	Programs with Site: Internal Medicine No.	%	Family Practice No.	%	Programs Planning to Have Site within Five Years: Internal Medicine No.	%	Family Practice No.	%
Inpatient geriatrics evaluation unit†	49	35	21	17	42	30	49	39
Inpatient geriatrics consult service	59	42	43	34	56	40	44	35
Geropsychiatry ward	38	27	23	18	24	17	26	21
Ambulatory geriatrics evaluation unit	48	34	51	40	40	29	41	33
Geriatrics clinic	55	39	41	33	37	26	39	31
Nursing home†	81	58	117	93	27	19	8	6

*To determine how well geriatrics has been integrated into residency training, the authors surveyed and obtained responses from 33% of all 378 family practice programs and all 420 internal medicine programs in the United States.

†Between specialties, $p < .01$.

Table 3

Average Numbers of Geriatrics Faculty Available and Needed in 140 Internal Medicine Programs and 126 Family Practice Programs That Either Had or Did Not Have a Geriatrics Curriculum, 1988*

	Average Number of Faculty			
	Internal Medicine Programs		Family Practice Programs	
Type of Faculty	With Curriculum ($n = 51$)	Without Curriculum ($n = 89$)	With Curriculum ($n = 101$)	Without Curriculum ($n = 25$)
Available				
Full-time geriatrics faculty‡	3.4	2.4	2.3	2.4
Full-time geriatrics faculty with formal training in geriatrics‡	1.1	0.3†	0.4	0.3
Part-time or voluntary geriatrics faculty	2.6	5.2	2.7	3.2
Needed				
Full-time geriatrics faculty currently needed‡	4.3	3.9	2.2	2.2
Full-time geriatrics faculty needed in five years‡	6.2	4.6	2.8	3.0

*To determine how well geriatrics has been integrated into residency training, the authors surveyed and obtained responses from 33% of all 378 family practice programs and all 420 internal medicine programs in the United States.

†$p < .01$ between the programs with curricula and those without curricula, within each specialty.

‡$p < .05$ between the internal medicine programs with curricula and the family practice programs with curricula.

ula in place had no formally trained faculty in geriatrics ($p < .01$) (not displayed in table). Programs with and without geriatrics curricula also differed with respect to available training sites. Inpatient GEUs and geriatrics consult services and ambulatory geriatrics clinics were common only in the IM programs that had geriatrics curricula (Table 4). In contrast, nursing homes were nearly universal in the FP programs that had geriatrics curricula.

More than 85% of geriatrics curricula in both specialties used an interdisciplinary approach (a team of physician and nonphysician health care providers with diverse areas of expertise), but internal medicine offered a greater number of interdisciplinary conferences per month (3.4 versus 2.1, $p < .001$). Nevertheless, residents in the two disciplines attended almost the same number of conferences during their residencies (averages 14.9 and 14.3, internal medicine and family practice, respectively). Research

Table 4

Percentages of 140 Internal Medicine Programs and 126 Family Practice Programs with Clinical Sites for Teaching Geriatrics and with Plans for Such Sites; Programs Categorized as Those with and Those without a Geriatrics Curriculum, 1988*

	Percentage of Programs with Site				Percentage of Programs Planning to Have Site within Five Years			
	Internal Medicine Programs		Family Practice Programs		Internal Medicine Programs		Family Practice Programs	
Type of Clinical Site	With Curriculum	Without Curriculum	With Curriculum	Without Curriculum	With Curriculum†	Without Curriculum†	With Curriculum†	Without Curriculum†
Inpatient geriatrics evaluation unit	63	22	21	9	22	37	21	9
Inpatient geriatrics consult service	73	25	39	22	22	51	37	35
Geropsychiatry ward	35	18	24	5	28	21	23	19
Ambulatory care geriatrics evaluation unit	64	21	47	22	23	35	33	36
Geriatrics clinic	81	19	40	17	17	35	33	36
Nursing home	80	49	96	78	12	24	3	22

*To determine how well geriatrics has been integrated into residency training, the authors surveyed and obtained responses from 33% of all 378 family practice programs and all 420 internal medicine programs in the United States.

†Curriculum status at the time of this study.

opportunities in geriatrics were available in approximately two-thirds of the programs that had geriatrics curricula in both specialties, but in the preceding two years few residents had participated in geriatrics research (average 1.1 in internal medicine and 0.45 in family practice, $p < .05$).

In response to our attitude questions, the FP respondents rated the importance of specific training in geriatric medicine higher than did the IM respondents ($p < .001$). Among the IM group, the geriatricians rated the importance of specific geriatrics training higher than did the department chairs and/or program directors ($p < .001$).

Overall, the most frequently cited major obstacle to implementing geriatrics curricula was the lack of faculty. This observation held true in every subgroup (curriculum in place, planned, and not planned) and in both specialties. The only exception was the FP programs that had a curriculum in place. The major obstacle for this subgroup was that the FP curriculum was "too full." Overall, a full curriculum was the second most frequently cited major obstacle. Resistance by residents and faculty, lack of training sites, and inadequate patient census were mentioned less frequently as major obstacles.

Discussion

How well integrated into residency training is geriatrics? Our results indicated that 80% of our sample of the U.S. family practice programs, but only 36% of our sample of the internal medicine programs had geriatrics curricula in place. If all the programs that were planning to institute geriatrics curricula within five years are able to achieve their goals, these figures will rise to 95% for family practice and 78% for internal medicine. Although this increase in geriatrics curricula from 21% in 1979[10] to our combined rate of 57% is laudable, the rate of growth at this level of training is less than that at other levels of medical education.[5] Furthermore, the discrepancy between family practice and internal medicine is of concern. In addition to having fewer geriatrics curricula in place, only 61% of the internal medicine curricula were required, compared with 92% of those in family practice.

Although more residency programs have instituted geriatrics curricula, these programs can be effective only if they have adequate faculty and training sites. Our data suggest that, in general, current and future faculty resources are inadequate to meet the needs of teaching geriatrics in residency training. Our respondents indicated that there is currently an average mismatch between existing and needed full-time-equivalent geriatrics faculty of 1.2 positions per program in internal medicine residency programs. If we extrapolate our findings, we calculate that in order to correct this degree of mismatch for all the 420 U.S. internal medicine programs, an additional 500 full-time-equivalent geriatrics faculty are needed. Using similar calculations for all U.S. programs based on the perceptions of needs in five years reported in the responses to our survey, another 500 full-time faculty will be needed in internal medicine and 225 will be needed in family medicine. This need for an additional 1,225 geriatrics faculty specifically for residency education and training in the next five years far exceeds the potential to train geriatrics faculty even if the Institute of Medicine is report's recommendation to double fellowship training capacity in geriatrics were adopted.[16] Moreover, even if training opportunities were available, the discipline is plagued by a dearth of qualified physician applicants for geriatrics fellowship training. Consistent with these findings, the major obstacle in implementing geriatrics curricula has been the lack of faculty. In spite of this lack of faculty, fewer than one-third of the institutions responding reported having specific programs in geriatrics faculty development. Clearly, the training of faculty in geriatric medicine is emerging as an important medical education and health policy priority.

Clinical training sites also appear to be insufficient and inadequate at present. The majority of the residency programs in both specialties had neither inpatient nor ambulatory geriatrics evaluation units available for training. These units have become the cornerstone for patient care and for educational efforts in geriatrics,[17] and they provide natural settings for teaching geriatrics using an interdisciplinary approach.[18] Similarly, the number of geropsychiatry units is insufficient. In contrast, nursing homes have rapidly been accepted as training sites. Although the use of institutional long-term care settings as teaching sites is to be commended, we are concerned that some programs may be regarding a nursing home clinical experience as sufficient exposure to geriatric medicine. In fact, institutional long-term care is only a small component of clinical geriatrics. The relative paucity of other available clinical sites, even among programs that currently have geriatrics curricula, supports this concern.

The second major obstacle identified in our study was the perception that the residency curriculum was too full to adequately develop a geriatrics curriculum. We believe that this obstacle can best be overcome by a reevaluation of the educational priorities of residency training. Given the current and projected proportions of patient care that family practitioners and internists devote to the primary care of older persons, can mandatory intensive care and emergency room rotations be justified at the expense of geriatrics rotations? Perhaps increased training in geriatrics can be coupled with other changes in the structure of residency training. For example, the Accreditation Council for Graduate Medical Education mandates that at least 25% of time be spent in ambulatory training in internal medicine programs[9] may be addressed by development rotations or longitudinal experiences in ambulatory geriatrics evaluation units.

The implementation of structural changes in residency programs to increase geriatrics content will require attitude changes, particularly in internal medicine. In general, internal medicine faculty completing our survey ranked the importance of geriatrics training lower than did the family practitioners. Furthermore, within internal medicine, program directors and/or departmental chairs ranked the importance of geriatrics lower than did geriatricians. Efforts to increase geriatrics training in residency programs will undoubtably require the support of program directors, departmental chairs, and deans. We also believe that the residency review committee requirement of a formal geriatrics curriculum for family practice[8] has greatly accelerated the incorporation and acceptance of geriatrics as an integral part of residency training in that specialty. Similar support from the Accreditation Council for Graduate Medical Education[9] will probably promote the universal development of geriatrics curricula in internal medicine residency programs.

In interpreting our results, the potential biases of survey research must be noted. To reduce selection bias, we randomly selected a large, stratified, national sample. Another potential major problem, responder bias, was eliminated by obtaining a 100% response rate.

What future research is necessary, given the results of this study? The next step in assessing geriatrics teaching in residency training will be to define curricular content and assess how well programs are meeting their specific training goals. The effects of these curricula on improving attitudes, clinical skills, and knowledge in caring for the elderly population also need to be more stringently examined. Finally, the effect of improved training in geriatrics on quality of care provided to elderly persons must be evaluated.

In summary, some progress has been made towards including geriatrics content in residency training. Nevertheless, curriculum implementation in geriatrics is still far from universal. Moreover, serious obstacles may limit the expansion of geriatrics training in residency. The most important of these appear to be a shortage of faculty, a lack of adequate training sites, and possibly a lack of support from department chairs and program directors. Until these obstacles are overcome, the ability of residency programs to provide satisfactory training for internists and family practitioners in the care of elderly persons remains in doubt.

This project was supported by The Merck Company Foundation and the California Geriatric Education Center. The work was completed while Dr. Reuben was a John A. Hartford Faculty Development Scholar. We thank Jan Chernoff and Susan Kelleher for their help in obtaining survey responses, and all the participants in the survey.

References

1. American Association of Retired Persons, and Administration on Aging, U.S. Department of Health and Human Services. *A Profile of Older Americans—1989.* Washington, D.C.: American Association of Retired Persons, 1989, p. 2.
2. National Institute on Aging. *Personnel for Health Needs of the Elderly through Year 2020.* Bethesda, Maryland, National Institute on Aging, 1987.
3. Akpom, C. A., and Mayer, S. A Survey of Geriatric Education in United States Medical Schools, *J. Med. Educ.* **53**(1978): 66–68.
4. Vivell, S., Robbins, A. S., Solomon, D. H., and Beck, J. C. Education in the Clinical Years: Delusion or Reality, *Bull. N.Y. Acad. Med.* **61**(1985):520–533.
5. Vivell, S., Solomon, D. H., and Beck, J. C. Medical Education Responds to the 20th Century's Success Story, *J. Am. Geriatr. Soc.* **35**(1987):1107–1115.
6. Reuben, David B., et al. Unpublished data.
7. Geriatric Medicine: A Statement from the Federated Council for Internal Medicine. *Ann. Intern. Med.* **95**(1981):372–376.
8. Accreditation Council for Graduate Medical Education. *Directory of Graduate Medical Education Programs, 1988–1989.* Chicago, Illinois: American Medical Association, 1988, p. 25
9. Accreditation Council for Graduate Medical Education. *Special Requirements for Residency Training Program in Internal Medicine.* Chicago, Illinois: Implementation date 10/1/89, 1988.
10. Robbins, A. S., Vivell, S., and Beck, J. C. A Study of Geriatric Training Programs in the United States. *J. Med. Educ.* **57**(1982):79–86.
11. Calkins, E. Residency Training in Geriatric Medicine, the Geriatric Medical Education Imperative, 1984. *Bull. N.Y. Acad. Med.* **61**(1985):534–546.
12. Cefalu, C. A., Gehringer, G. R., and Mcknight, G. T. Geriatric Training in Family Practice. *J. Fam. Pract.* **11**(1980):971–972.
13. Goodenough G. K. Geriatric Teaching in Family Practice Residencies. *J. Fam. Pract.* **21**(1985):226–227.
14. Wallace, E. Association of Program Directors in Internal Medicine. Personal communication.
15. Sudman, S., and Bradburn, N. Improving Mailed Questionnaire Design. In D. C. Lockhard, ed., *Making Effective Use of Mailed Questionnaires.* New Directions for Program Evaluation, no. 21. San Francisco: Jossey-Bass, March 1984.
16. Rowe, J. W., Grossman, E., and Bond, E. and the Institute of Medicine Committees on Leadership for Academic Geriatric Medicine. Academic Geriatrics for the Year 2000. *N. Engl. J. Med.* **316**(1987): 1425–1428.
17. Epstein, A. M., et al. The Emergence of Geriatric Assessment Units. *Ann. Intern. Med.* **106**(1987):299–303.
18. Besdine, R. W. The Educational Utility of Comprehensive Functional Assessment in the Elderly. *J. Am. Geriatr. Soc.* **31**(1983):651–656.

Faculty Development and Organizational Systems Behavior

CHARLES E. HENLEY, D.O., M.P.H., and DAVID MAGELSSEN, M.D., M.A.

Abstract—Faculty development is that process that fosters improvement in faculty members' skills in teaching and research and promotes their career advancement. This study investigated the association between organizational behavior in military medical centers and the faculty development of its medical corps officers assigned to teaching positions. Such organizational behaviors as defining tasks clearly and resolving conflicts satisfactorily correlated well with the faculty members' overall satisfaction and other parameters of good faculty development. The results suggest that a strong relationship exists between the organizational behavior of an institution and the sense of identity, productivity, and continued career growth of its individual faculty members. *Acad. Med.* 65(1990):406–409.

This study examines the relationship between faculty and the organizational systems within which these faculty teach and work. Although the data were collected within the structure of the military medical center system, the authors maintain that the concepts can be generalized and applied to faculty development and organizational systems in other settings. Indeed, the pressures and requirements of faculty in terms of patient care, teaching, publishing, peer review, board certification, and other activities are essentially the same whether one is at a military medical center or at a comparable civilian university hospital. The same could be said for the essentials of the corresponding bureaucracies, in that the organizational behaviors of our teaching institutions determine the effectiveness and career satisfaction of their faculties and therefore the career development of the individual faculty member.

The military medical system in which this study was done is currently the largest single health care provider in the United States. At just one of its moderate-sized teaching centers, for example, there are over 100,000 outpatient visits each month. The challenge for military medicine is to procure and maintain a highly trained and skilled cadre of physician manpower to care for the ever-expanding pool of eligible patients. Since World War II, the backbone of this effort has been the military's postgraduate training programs, primarily situated in its modern centers and in selected community hospitals that have evolved as teaching centers.

Since the maintenance of these physician training programs is vital to the integrity and continuity of a modern medical corps, the military has assumed the responsibility of securing and developing the necessary faculty to staff each of these teaching programs. The challenge of faculty development requires that the careers of these faculty be thoughtfully developed to insure that the individuals remain productive and enthusiastic in their teaching disciplines and that the high quality of the training programs be maintained. In this regard, the institution's organizational behavior has a direct and significant impact on the ability to retain good faculty.

In attempting to investigate the organizational behavioral elements that have an impact on faculty development, the authors have used the military training programs as a microcosm within which the broader arena of academic medicine can be viewed. The combined faculty of the major teaching medical centers within the military were surveyed for this study. Within these institutions, key organizational elements were identified and studied as to their influences on faculty career development. Productive and nonproductive organizational behaviors were analyzed with successful faculty development as the measured outcome.

One of the critical questions in starting a study of this magnitude is first to clearly conceptualize what is meant by "faculty" and what is meant by "faculty development."

Historically, the teaching of most professional and academic subjects has been left to experts, who are usually quite knowledgeable in their fields, but often have no formal training as educators. The modern concept of faculty implies some skill beyond patient care that enables a person to transfer knowledge to a learner. This process has usually relied heavily on inspiration, hard work, and enthusiasm, but in the climate of academic medicine today, such skills as research methodology, program development and evaluation, and management skills, as well as knowledge of a variety of teaching methods, are complementary to the career of an academic physician.[1,2]

Faculty development, then, is that process that fosters improvement in faculty members' teaching and research skills and promotes their career advancement; a faculty development program should encompass all activities that assist faculty in these ways. Jason and Westberg[3] recommend a six-step strategy in creating a faculty development program:

1. Determine the expectations, needs, and goals of teachers.

Dr. Henley is program director of the Center's Faculty Development Fellowship Program, and clinical assistant professor, Department of Family Medicine, University of Washington School of Medicine, Seattle. At the time of this study, Dr. Magelssen was a fellow, Faculty Development Fellowship Program, Department of Family Medicine, Madigan Army Medical Center, Tacoma, Washington; he is now director, flexible intern program, and chief of benign gynecological surgery, Department of Obstetrics-Gynecology, both at the Center.

The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army or the Department of Defense.

Correspondence and requests for reprints should be addressed to Dr. Magelssen, Department of Obstetrics and Gynecology, Madigan Army Medical Center, Tacoma, WA 98431.

2. Share goals, expectations and convictions. Provide a forum for communication.

3. Jointly formulate goals in cooperation with participating faculty.

4. Jointly decide on strategies for meeting these goals. This might involve planning workshops, retreats, presentations, or demonstrations, or experimenting with theoretical models of instructional methodology.

5. Jointly decide on strategies for evaluating the teachers' progress. This often is the step that can potentially produce the greatest rewards or, conversely, the greatest disappointments in a program.

6. Implement the program. A program needs to have clear, realistic, obtainable goals.

The systematic adherence to these six steps can provide a framework from which to develop a faculty development program within an organization or evaluate an existing program. Using this strategy as a model, an analysis of the organizational system's influence upon the effectiveness of a faculty development program can be made.

Method

For the present study, the previously stated concepts of faculty development were objectified into the dimensions of personal esteem, professional career growth, intrinsic motivation, identity with task performance, group identity, realization of personal objectives, career satisfaction, and productivity. These eight variables of faculty development were compared with elements constituting organizational behavior, which were extracted from the literature[4,5] and consisted of systems structure, group norms, resource allocations, statement of mission, task definition and objectives, reward system, power and authority, information processing, group performance, and intragroup conflicts.

These dimensions of organizational behavior and faculty development were then codified into survey questions and formatted with a five-point Likert-type scale of strongly disagree, disagree, no opinion, agree, and strongly agree. There were 32 questions for both the faculty development and the organizational behavior dimensions, with between three and five questions describing each. There were also demographic questions dealing with the institution, department assignment, formal rank (for economic and tenure data), career longevity, and professional responsibility, for a total of 77 questions. The questionnaire as well as the key for which questions collate with each of the dimensions is available upon request.

In June 1988, questionnaires were distributed to 13 military medical teaching hospitals.* An estimated 2,331 teaching faculty were identified and surveyed at these institutions.

Results

A total of 1,097 surveys were returned by the given completion deadline. Two of the three Air Force hospitals surveyed initially declined, but subsequently conceded to participate at a later date after receiving approval from higher authority, although not in time for completion of this study. This accounted a rather low return rate from the Air Force. The overall return rate from all faculty was 47%.

Demographic Data

Approximately 98% of the faculty who responded had completed residency training, with 71% having done postgraduate work in a military program and the rest in a civilian institution.

Only 9% of the responding faculty stated that their hospitals or departments had existing faculty development programs, while 91% stated negatively or were not sure.

Of the 1,097 physicians completing the questionnaire survey, 99 were women and 998 were men.

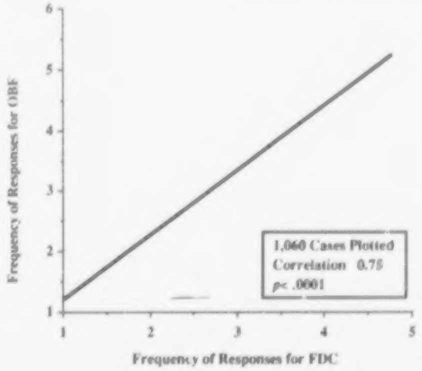


Figure 1. **Graphic representation of regression data (Pearson's *r*) taken from a scatter plot of the frequency of mean Likert scale responses for the Organizational Behavior Factor (OBF) compared with the frequency of the mean responses for the Faculty Development Coefficient (FDC). The strong linear relationship of the data reflects the correlation between the values for the OBF and the corresponding values for the FDC for the military teaching hospitals studied. As the favorable institutional behaviors such as clear goals and performance objectives, or good resolution of group conflicts, increased in the institutions studied, so also did those factors recognized as indicators of good faculty development, such as the faculty members' increasing motivation, productivity, and sense of self-esteem.**

---

*Army = Walter Reed Amy Medical Center, Washington, D.C.; William Beaumont Army Medical Center, El Paso, Texas; Brooke Army Medical Center, San Antonio, Texas; Letterman Army Medical Center, San Francisco, California; Fitzsimmons Army Medical Center, Denver, Colorado; Tripler Army Medical Center, Honolulu, Hawaii; Madigan Army Medical Center, Tacoma, Washington; and Dwight David Eisenhower Army Medical Center, Augusta, Georgia. Navy = Bethesda Naval Hospital, Bethesda, Maryland; Portsmouth Naval Hospital, Portsmouth, Virginia; Oakland Naval Hospital, Oakland, California; and Balboa Naval Hospital, San Diego, California. Air Force = Keesler AFB Regional Medical Center, Biloxi, Mississippi.

Table 1

Relative Degrees of Satisfaction with Organizational and Faculty Development Dimensions Shown by 1,097 Faculty at 13 Military Teaching Hospitals by the Ranking of Their Mean Scores for These Dimensions, 1988*

Dimension	Rank of Mean Scores of Faculty from Each Branch of Military Service†		
	Army	Navy	Air Force
Organizational			
Task objectives	10th	10th	10th
Group performance	9th	8th	9th
Group norms	8th	7th	6th
Conflict resolution	7th	9th	7th
Clarity of mission	6th	6th	8th
Information processing	5th	5th‡	5th
Authority and power	4th	4th‡	4th‡
Systems structure	3rd‡	2nd‡	3rd‡
Resources	2nd‡	1st‡	2nd‡
Reward system	1st‡	3rd‡	1st‡
Faculty development			
Motivation	8th	8th	8th
Self esteem	7th	7th	7th
Productivity	6th	6th	6th
Task identity	5th	5th	5th
Group identity	4th	4th	4th
Career growth	3rd	3rd	3rd
Career satisfaction	2nd	2nd	2nd
Personal objectives	1st	1st	1st‡

*The authors distributed a questionnaire to an estimated 2,331 teaching faculty at 13 military teaching hospitals; 1,097 faculty completed the questionnaire satisfactorily for use in this study. The questionnaire presented questions concerning various dimensions of faculty development and organizational behavior, and the faculty rated these dimensions on a five-point Likert-type scale, where 1 = least satisfactory and 5 = most satisfactory. This table shows the hierarchy of the mean scores from each service branch's hospitals.

†A higher number indicates a higher ranking of a mean score. For example, among the ten organizational dimensions, task objectives received the highest rank (10) from the faculty in each branch, indicating that, relative to the other organizational dimensions, the faculty felt that their daily work expectations were generally understood. Among the eight faculty development dimensions, motivations received the highest rank, indicating that, relative to the other faculty development dimensions, the faculty felt that their motivation was high.

‡The mean score indicates that fewer than 50% of the responding faculty felt that this dimension was satisfactory.

Survey Results

A composite score was tabulated from the mean of the total scores of the survey questions that related to either organizational behavior or faculty development. These values were termed the Organizational Behavior Factor (OBF), and the Faculty Development Coefficient (FDC), respectively. The OBF and the FDC were calculated for 1,097 respondents; however, 37 questionnaires used for the comparison between the OBF and the FDC had to be discarded due to missing values, leaving a total of 1,060 eligible questionnaires.

A mean value for each of the different elements was determined from the Likert scores for each of the questions pertaining to that element, and a composite mean score was then determined for both the OBF and the FDC. The mean value for the OBF was 3.587, with a median score of 3.623 and a mode of 3.844, for 1,060 cases. The mean for the Faculty Development Coefficient (FDC) was 2.973, with a median of 3.0 and a mode of 3.063, for 1,097 cases.

When the mean scores for OBF and the FDC for 1,060 respondents were compared using Pearson's product moment correlation coefficient, a strong linear association was found ($p < .0001$) (Fig. 1). This would imply that as an increase in supportive organizational behavior was perceived among the faculty, the perception of their own faculty career development improved as well.

The Army faculty had significantly higher scores for the OBF than did the faculties from the Navy and Air Force when the individual elements of organizational behavior were compared. There was, however, a remarkable consistency between all groups in terms of the ranking and perceived satisfaction with each of the ten elements of organizational behavior (Table 1). Elements of organizational behavior that directly affect daily functioning received the highest ranking across the board. Faculty gave the highest mean score to task objectives, indicating that they felt their daily work expectations were generally understood. Other local group behaviors, such as overall work performances, group norms, and resolution of group conflicts, also had higher mean scores.

Organizational behavior that was centered in the less personal, larger systems structure received the lowest scores. A mean score of 3 or less for any OBF element occurred where the faculty perceived this organizational behavior to be unsupportive. The overall medical department structure, the availability of resources, and the formal reward system had the lowest mean scores from all faculty.

Despite the poor scores for systems behavior, reward systems, and resources, the faculty members surveyed consistently viewed themselves as being productive and highly motivated and possessing strong self-esteem (Table 1). Overall, the mean scores for the faculty development elements were higher than were those for the organizational behavior elements, with only three faculty development elements having mean scores less than 3.0. Faculty development factors that reflected more personal control by the respondents, namely intrinsic motivation, self-esteem, and productivity received the highest mean scores from all groups. Those factors that were more dependent upon the response and support of the organizational structures, such as career growth, career satisfaction, and meeting personal objectives, had the lowest mean scores.

Summary and Conclusions

A strong linear relationship was demonstrated between organizational behavior and faculty development in 13 teaching medical centers within the military medical system.

Although most of the responding faculty perceived themselves as being highly motivated and productive in their professional activities, all groups gave their career growth and the opportunity to meet their personal objectives some of the lowest ratings of the faculty development dimensions listed in the questionnaire. The greatest obstacles the responding faculty perceived to good faculty development were a lack of resources and a poor reward system.

While these data were obtained using faculty from military teaching programs, the authors maintain that the findings would have the same meaning, and deliver the same message, if they had been obtained from faculty of nonmilitary academic institutions, such as university teaching faculty. In either kind of institution, a reward system that is not valued or does not appropriately recognize achievement by faculty members will ultimately be counterproductive to that institution. In the same vein, a system that is perceived as sluggish, complex, and unresponsive in procuring adequate resources for its members will impair morale and productivity.

The implications from these data are that faculty retention, job satisfaction, and career development are areas that can be enhanced if proper attention is paid to certain organizational characteristics. Supporting junior faculty practices and research within the institution by streamlining the procurement and grants process, instituting a reward system that reflects the talents and contributions of the individual and that is valued by the faculty members, as well as generally personalizing the organizational behavior of a particular department or institution, may very well be examples of how a program can produce a more supportive, creative environment and promote greater productivity among its faculty. The creation of a formal faculty development program within a department of institution is an organizational behavior that should be considered as one that would help faculty gain the insight and skills required to be more effective in their work, which would in turn enhance the progress of their careers. Of interest, and a potential area of future studies, are the relationships between faculty development and organizational behaviors among certain subgroups such as women and other minorities, and specific medical and surgical subspecialties.

References

1. Pickering, George. *Quest for Excellence in Medical Education.* Oxford, England: Oxford University Press, 1978.
2. Martin, L. F. The Importance of Management Training for Academic Physicians. *Surgery* **95**(1985):52–54.
3. Jason, Hilliard, and Westberg, Jane. *Teachers and Teaching in U.S. Medical Schools.* Hartford, Connecticut: Appleton-Century-Crofts, 1982.
4. Gordon, Judith R. *A Diagnostic Approach to Organizational Behavior.* Newton, Massachusetts: Allyn and Bacon, 1983.
5. Lichtenstein, Richard. Measuring the Job Satisfaction of Physicians in Organized Settings. *Med. Care* **22**(1983):56–68.

AAMC DATA REPORT

Results of the NRMP for 1990

The number of U.S. seniors who enrolled in the National Resident Matching Program (NRMP) in 1990 decreased to 15,254 from 15,438 in 1989. The number who subsequently withdrew increased by 25 (Table 1), and the number of active applicants decreased by 209. Matched U.S. seniors totaled 12,976 (93.3%) in 1990, compared with 13,215 (93.6) in 1989.

The total number of other U.S. and Canadian citizens ("fifth pathway" students, physician applicants [earlier U.S. graduates seeking graduate training], osteopathic and Canadian graduates) who applied to the Match this year decreased by 238 to 3,345. Among the other U.S. and Canadian citizens, the number of "fifth pathway" applicants decreased from 171 to 114; 82 participated in the Match, and 72% were successful. The number of graduates of osteopathic schools who initially applied to the Match decreased by seven percentage points to 1,188; 60.7% were matched. The number of physician applicants decreased by five percentage points, and 50% were successful in the Match.

The number of U.S. citizen-FMGs who applied decreased by 209 to 1,306; 55.6% successfully matched, an increase of 2.3 percentage points. The number of foreign-born–FMGs who initially enrolled decreased by 160 (4%) to 3,935; however, 1,044, compared with 1,038 in 1989, were active at the time of the Match. The number who matched increased by 5.5 percentage points to 59.7%.

The number of positions offered for the first postgraduate year (PGY1) broke the 20,000 mark in 1990, when 20,101 positions were offered. The numbers offered in family practice, three-year programs in internal medicine, and pediatrics decreased. Increases occurred in one–year internal medicine programs, obstetrics and gynecology, anesthesiology, emergency medicine, and diagnostic radiology (Table 2). The ratio of positions per U.S. senior student initially enrolled was 1.3:1 in 1990.

The overall filling of PGY1 positions decreased by one percentage point in 1990 to 79.2%, and the filling of positions by U.S. seniors decreased by 1.6 percentage points. Since 1985, filling by U.S. seniors has decreased by nearly ten percentage points and total filling by just under nine. The most competitive programs were those in emergency medicine, obstetrics and gynecology, orthopedic surgery, general surgery, and diagnostic radiology, where over 90% of positions were filled with U.S. seniors. Family practice, internal medicine, pediatrics, psychiatry, and pathology filled at or below the 64.6% average with U.S. seniors.

An increased number of Advanced Programs for Students (PGY2) positions were offered in the Match this year. Applicants matching to these positions will do a year of broad clinical training—generally in internal medicine, general surgery, or transitional year—before entering advanced positions in 1991. Three hundred twenty-six programs offered 1,656 positions, which were 75.1% filled by 1,244 U.S. seniors and 84.6% filled by 1,401 total applicants. One hundred thirteen programs in diagnostic radiology offered 429 PGY2 positions, of which 80.9% were filled with U.S. seniors and 93% were filled overall. One hundred eleven programs in anesthesiology offered 822 positions, of which 642 (78%) were filled with U.S. senior students and 41 (6%) were filled with other applicants, for a total filling of 84.3%. Other advanced programs with significant applicant matches included those in physical medicine (102 matched, 76.7%), emergency medicine (74 matched, 60.7%), orthopedic surgery (43 matched, 87.8%), and psychiatry (35 matched, 38.9%).

In the 1990 Match, 794 applicants participated as 397 couples; 365 of these couples were U.S. seniors. Ninety-seven percent of the couples were matched successfully.

Data describing the trends in positions offered and filled by U.S. seniors and other applicants during the past several years are published annually in April in *NRMP Data*. It is available upon written request to the NRMP, One Rotary Center, Suite 807, Evanston, IL 60201.

This report was prepared by the National Resident Matching Program.

Table 1

Participation in the National Resident Matching Program in 1990

	U.S. Senior Students		Other U.S. and Canadian Students*		U.S. Foreign Graduates		Foreign-born Foreign Graduates		Total Applicants	
Category	No.	%	No.	%	No.	%	No.	%	No.	%
Total applicants	15,254	100.0	3,345	100.0	1,306	100.0	3,935	100.0	23,840	100.0
Withdrew	1,346	8.8	1,407	42.1	477	36.5	1,606	40.8	4,836	20.3
Active participants	13,908	91.2	1,938	57.9	829	63.5	2,329	59.2	19,004	79.7
Matched for PGY1	12,976	93.3	1,085	56.0	461	55.6	1,390	59.7	15,912	83.7
Unmatched for PGY1	932	6.7	853	44.0	368	44.4	939	40.3	3,092	16.3
Total	13,908	100.0	1,938	100.0	829	100.0	2,329	100.0	19,004	100.0

*Other U.S. and Canadian students consist of graduates from U.S. osteopathic schools and Canadian medical schools; physician candidates, that is, earlier U.S. graduates seeking graduate training positions; "fifth-pathway" students, who are U.S. nationals who have completed course requirements for a medical degree abroad followed by a year of clinical training in a U.S. hospital under medical school sponsorship.

Table 2

Positions Offered and Filled by Types of Applicants in 1990

	PGY1 Positions Offered		Applicants Matched											
			U.S. Senior Students			Other U.S. Applicants‡			Foreign-born Foreign Graduates			Total Applicants		
				%			%			%			%	
Specialty	No.	%*	No.	*	†	No.	*	†	No.	*	†	No.	*	†
Family practice	2,393	11.9	1,418	10.9	59.3	200	12.9	8.4	67	4.8	2.8	1,685	10.6	70.4
Internal medicine	7,442	37.0	4,569	35.2	61.4	512	33.1	6.9	825	59.4	11.1	5,906	37.1	79.4
Pediatrics	2,052	10.2	1,286	9.9	62.7	156	10.1	7.6	227	16.3	11.1	1,669	10.5	81.3
Obstetrics-gynecology	1,076	5.4	925	7.1	86.0	94	6.1	8.7	25	1.8	2.3	1,044	6.6	97.0
Psychiatry	1,135	5.6	664	5.1	58.5	124	8.0	10.9	75	5.4	6.6	863	5.4	76.0
Medical specialties	81	0.4	38	0.3	46.9	6	0.4	7.4	9	0.6	11.1	53	0.3	65.4
Dermatology	15	0.1	15	0.0	100.0	0	0.0	—	0	0.0	—	15	0.0	100.0
Neurology	51	0.3	15	0.1	29.4	5	0.3	9.8	8	0.6	15.7	28	0.2	54.9
Ophthalmology	15	0.1	8	0.1	53.3	1	0.1	—	1	0.1	—	10	0.1	66.7
General surgery	2,255	11.2	1,501	11.6	66.6	133	8.6	5.9	74	5.3	3.3	1,708	10.7	75.7
Surgical specialties	633	3.1	539	4.2	85.2	62	4.0	9.8	4	0.3	0.6	605	3.8	95.6
Neurosurgery	38	0.2	25	0.2	65.8	3	0.2	7.9	0	0.0	—	28	0.2	73.7
Orthopedic surgery	456	2.3	400	3.1	87.7	52	3.4	11.4	1	0.1	—	453	2.8	99.3
Otolaryngology	52	0.3	42	0.3	80.8	3	0.2	5.8	2	0.1	—	47	0.3	90.4
Urology	79	0.4	65	0.5	82.3	4	0.3	5.1	0	0.0	—	69	0.4	87.3
Plastic surgery	8	0.0	7	0.0	87.5	0	0.0	—	1	0.1	—	8	0.0	100.0
Support specialties	1,706	8.5	1,172	9.0	68.7	187	12.1	11.0	34	2.4	2.0	1,395	8.8	81.8
Anesthesiology	321	1.6	251	1.9	78.2	9	0.6	2.8	4	0.3	1.2	264	1.7	82.2
Emergency medicine	440	2.2	337	2.6	76.6	71	4.6	16.1	2	0.1	0.5	410	2.6	93.2
Pathology	486	2.4	243	1.9	50.0	24	1.6	4.9	25	1.8	5.1	292	1.8	60.1
Physical medicine	95	0.5	71	0.5	74.7	7	0.5	7.4	0	0.0	0.0	78	0.5	82.1
Preventive medicine	15	0.1	5	0.0	33.3	0	0.0	—	1	0.1	6.7	6	0.0	40.0
Diagnostic radiology	337	1.7	258	2.0	76.6	76	4.9	22.6	2	0.1	0.6	336	2.1	99.7
Therapeutic radiology	6	0.0	5	0.0	83.3	0	0.0	—	0	0.0	0.0	5	0.0	83.3
Nuclear Medicine	6	0.0	2	0.0	33.3	2	0.1	33.3	0	0.0	0.0	4	0.0	66.7
Transitional	1,328	6.6	864	6.7	65.1	70	4.5	5.3	50	3.6	3.8	984	6.2	74.1
Total	20,101	100.0	12,976	100.0	64.6	1,546	100.0	7.7	1,390	100.0	6.9	15,912	100.0	79.2

*Percent distributions of positions offered and filled among specialties (percentages may not add to 100 due to rounding).
†Percent of positions offered in specialty that were filled by applicants.
‡Students listed in the footnote to Table 1 plus U.S. foreign medical graduates.

A3273-8

JULY 1990

ACADEMIC MEDICINE

JOURNAL OF THE ASSOCIATION OF AMERICAN MEDICAL COLLEGES

- HMOs and Medical Education
- Computers in Clinical Education
- Theme: AIDS and Attitudes
  - Medical, Dental, and Nursing Students
  - Impact of AIDS Information
  - Willingness to Treat AIDS Patients
  - Faculty and Students
- In-Training Examinations
- Effect of Minority Enrichment Program
- AAMC Data Report: Debt and Specialties
- AAMC Paper: Conflict of Interest

VOLUME 65 • NUMBER SEVEN

JULY 1990
VOLUME 65
NUMBER SEVEN

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Special Article: Health Maintenance Organizations and Medical Education: Breaking the Barriers, *by Gordon* T. *Moore*, M.D. 427

Ideas for Medical Education Feedback in an Objective Structured Clinical Examination by Medical Students Serving as Patients, Examiners, and Teachers, *by Ilene* B. *Harris*, *Ph.D.*, *and Wesley* J. *Miller*, M.D. 433

Medical Informatics New Thoughts about Medical Students as Effective Searchers of MEDLINE, *by Joyce* A. *Mitchell*, *Ph.D.*, E. *Diane Johnson*, M.L.S., *and Virginia* K. *Proud*, M.D. 434

Academic Computing in Medicine *Preface*: Documenting Clinical Educational Experiences, *by Robert* F. *Rubeck*, *Ph.D.* 438

Using a Computerized Database to Manage a Decentralized Ambulatory Care Clerkship, *by Thomas* J. *Ruane*, M.D., *and Blake* W. H. *Smith*, *Ph.D.* 438

Developing a Computer-based System to Document and Report Students' Clinical Patient Encounters, *by Donald* B. *Witzke*, *Ph.D.*, *Nancy Alexander Koff*, *Ph.D.*, *Anna* M. *McGeagh*, *M.Ed.*, *and Patricia Dean Skinner*, M.S. 440

From the President 442

AAMC Focus 443

National Policy Perspectives 444

Commentary AIDS: Just Another Disease? *by Louis* J. *Kettel*, M.D. 446

Research Reports
ARTICLES A Survey of Medical School Teachers to Identify Basic Biomedical Concepts Medical Students Should Understand, *by Beth Dawson-Saunders*, *Ph.D.*, *Paul* J. *Feltovich*, *Ph.D.*, *Richard* L. *Coulson*, *Ph.D.*, *and David* E. *Steward*, M.D. 448

Predictive Validity of the American Board of Family Practice In-Training Examination, *by Terrence* M. *Leigh*, *Ed.D.*, *Timothy* P. *Johnson*, *Ph.D.*, *and Nicholas* J. *Pisacano*, M.D. 454

SHORT PAPERS

Medical and Dental Students' Attitudes about the AIDS Epidemic, *by Carol* A. *Bernstein*, M.D., *Judith* G. *Rabkin*, Ph.D., *and Henrietta Wolland*, M.P.H. 458

The Impact of an AIDS Symposium on Attitudes of Providers of Pediatric Health Care, *by Lloyd* R. *Feit*, M.D., *Sanford* M. *Melzer*, M.D., *Sten* H. *Vermund*, M.D., M.Sc., *and Steven* P. *Shelov*, M.D. 461

Attitudes of Medical School Faculty and Students toward Acquired Immunodeficiency Syndrome, *by Theodore* B. *Feldmann*, M.D., *Roger* A. *Bell*, Ed.D., *Judith* J. *Stephenson*, S.M., *and Frances* E. *Purifoy*, Ph.D. 464

Medical and Nursing Students' Attitudes about AIDS Issues, *by Bonnie* J. *Tesch*, M.D., *Deborah* E. *Simpson*, Ph.D., *and Barbara* D. *Kirby*, M.D. 467

Predictors of Medical and Nursing Students' Levels of HIV-AIDS Knowledge and Their Resistance to Working with AIDS Patients, *by Thomas* J. *Ficarrotto*, Ph.D., *Margaret Grade*, Ph.D., *Nancy Bliwise*, Ph.D., *and Thomas Irish*, Ph.D. 470

Willingness of Health-Professions Students to Treat Patients with AIDS, *by Charles* J. *Currey*, M.H.A., P.A.-C., *Michael Johnson*, M.D., *and Barbara Ogden*, M.S.N. 472

Analysis of Residents' Performances on the In-Training Examination of the American Board of Anesthesiology–American Society of Anesthesiologists, *by James* R. *Hall*, M.D., *and George* A. *Cotsonis*, M.A. 475

Minority Students' Career Choices and Education Five Years after They Completed a Summer Enrichment Program, *by Vera* B. *Thurmond*, Ed.D., *and Antonio Mott*, M.D. 478

The Relationship of Minority Students' MCAT Scores and Grade Point Averages to Their Acceptance into Medical School, *by Kathleen* B. *Lynch*, Ph.D., *and Moses* K. *Woode*, Ph.D. 480

Brief Communications 482

AAMC Data Report Graduating Medical Students' Debt and Specialty Choices, *by Randolph Park*, Ph.D. 485

AAMC Papers Guidelines for Dealing with Faculty Conflicts of Commitment and Conflicts of Interest in Research 487

JULY 1990
VOLUME 65
NUMBER SEVEN

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
Edward J. Stemmler, M.D.
  Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
  St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
  Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
  Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
  University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
  University of Kentucky College of Medicine

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
  Georgetown University Hospital
Philip Anderson, M.D.
  University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
  Harvard Medical School
Eta Berner, Ed.D.
  University of Alabama School of Medicine
Ruth Bulger, Ph.D.
  National Academy of Sciences
Vincent A. Fulginiti, M.D.
  Tulane University School of Medicine
Nancy Gary, M.D.
  Albany Medical College
Beatrix Hamburg, M.D.
  Mt. Sinai Hospital
Murray M. Kappelman, M.D.
  The University of Maryland School of Medicine
Ralph W. Muller
  University of Chicago Hospitals
Herbert Pardes, M.D.
  College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
  University of Alabama School of Medicine
W. Vickery Stoughton
  Toronto General Hospital
Martha Vaughan, M.D.
  National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
David H. Cohen, Ph.D.
  Northwestern University

Chairman-Elect
William T. Butler, M.D.
  Baylor College of Medicine

Immediate Past Chairman
D. Kay Clawson, M.D.
  University of Kansas Medical Center School of Medicine

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
  George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Joe Dan Coulter, Ph.D.
  University of Iowa College of Medicine

Chairman, Council of Teaching Hospitals
Raymond G. Schultze, M.D.
  University of California, Los Angeles Medical Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.

Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036; (202) 828-0416.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

RESEARCH REPORTS

ARTICLES

A Survey of Medical School Teachers to Identify Basic Biomedical Concepts Medical Students Should Understand

BETH DAWSON-SAUNDERS, Ph.D., PAUL J. FELTOVICH, Ph.D.,
RICHARD L. COULSON, Ph.D., and DAVID E. STEWARD, M.D.

Abstract—Insights from the cognitive sciences indicate a continuing need for physicians to understand conceptual knowledge from the basic sciences, despite recent concerns regarding the increasing amount of information in medicine and the growing emphasis on performance skills. A 1987 survey of selected basic science and clinical teachers in North American medical schools was undertaken to identify basic biomedical concepts that are important in the practice of medicine and to specify how difficult these are for students to learn, apply, or both. Responses from faculty (nominated by their deans to answer the survey) from 82% of the medical schools indicated considerable agreement between the basic science teachers and clinical teachers on the relative importances of a set of biomedical concepts, and showed relatively minor levels of disagreement on how difficult these concepts are. The judgments of these teachers could prove extremely useful in (1) determining concepts that—because of their importance—should receive special attention in curriculum efforts, and (2) determining concepts that—because of their difficulty—need "special handling." *Acad. Med.* 65(1990):448-454.

At least two recent trends have converged to create what can be seen as a confusing picture within medical education regarding the role of conceptual knowledge from the basic sciences in the education of the physician. The first trend is what is commonly referred to as the "information explosion," the proliferation of new and changing knowledge in fields pertinent to medical practice. One response to the proliferation of medical information has involved proposals for better technologies for providing access to information.[1] Another response addresses the curriculum for medical education; if the knowledge base is constantly changing and expanding, then it is not feasible to try to teach it all. This response has led, for instance, to an emphasis on teaching the skills of knowledge acquisition or self-directed learning and to calls for greater focus within the curriculum.[2,3] The second trend involves uncertainty about the relationship between conceptual knowledge of subject matter and the professional tasks and performances of the physician.[4] Perhaps influenced also in part by a dissatisfaction with standard methods of testing for knowledge, this trend has engendered an increased emphasis on performance, both in the training for professional skills and in the assessment of professional competence.[2,3,5]

Both of these trends could lead medical education toward a de-emphasis of basic conceptual knowledge from pertinent subject areas (such as the basic biomedical sciences), in favor of skills of performance and knowledge acquisition. However, some recent developments in the cognitive sciences should give pause to any movement of this sort. The first is the indication that conceptual knowledge and skills of performance and learning are highly intertwined and build off each other, so that neither can be slighted in the acquisition of competence.[6,7] The second is evidence that suggests that expert problem solvers are sometimes guided in their performances by basic concepts and principles within the subject matter of their fields.[8,9] The third involves mounting evidence that some difficult scientific concepts are rarely understood well by students and that misconceptions, once acquired, are difficult to correct.[10,11]

These three cautions have been reinforced in the reappraisal of the GPEP report (that is, the report of the Panel on the Geneval Professional Education of the Physician and College Preparation for Medicine) conducted by the Council of Deans and the Council of Academic Societies working groups of the Association of American Medical Colleges.[12] Their recommendations can be inter-

---

Dr. Dawson-Saunders is senior psychometrician, Division of Psychometrics, National Board of Medical Examiners, Philadelphia, Pennsylvania. Dr. Feltovich is associate professor, Department of Medical Education; Dr. Coulson is professor, Department of Physiology; and Dr. Steward is associate professor, Department of Internal Medicine; all three are at Southern Illinois University School of Medicine, Carbondale and Springfield, Illinois.

Correspondence and requests for reprints should be addressed to Dr. Dawson-Saunders, National Board of Medical Examiners, 3930 Chestnut St., Philadelphia, PA 19104.

preted as acknowledging that medical information and medical concepts are not necessarily the same, and that some basic medical concepts are important and must be understood well. In addition, it should be recognized that an emphasis on fostering clinical performance skills and learning skills can be compatible with an emphasis on understanding biomedical concepts.[13] But the information explosion still exists, and even the realm of biomedical science concepts is large. If the curriculum is to prune some teaching of information in favor of fostering deep understanding of concepts — and assuming there must be some focus even within concepts — how is this focus to be guided?

The present paper reports results of a 1987 survey of medical school teachers (both basic science and clinical) that was designed to identify basic biomedical science concepts that these teachers considered to be both important for achieving competence in the practice of medicine and difficult for students to learn, understand, and apply. (The survey was one component of a larger research project investigating biomedical understanding and the role of biomedical science knowledge in the clinical reasoning and performance of the physician.) In this article, the development and nature of the survey are presented first, followed by the results and their implications for medical school curricula and research.

Method

The design of the survey and analysis of the results required several steps: defining the content of the questionnaire, designing its format, the two-stage sampling plan, and the analysis.

Choosing Survey Concepts

A set of biomedical concepts that could be included in the survey was identified through interviews conducted with the teaching faculty at Southern Illinois University School of Medicine, Carbondale and Springfield, and with basic science researchers and teachers at a national physiology meeting. The disciplines represented included five basic sciences and five different areas of internal medicine. The persons interviewed were asked to nominate and discuss those basic science concepts "important to the practice of medicine" that their teaching had shown were difficult for students to understand and apply. They were also asked to identify important concepts that, although complex, they judged to be less difficult. Content descriptions were obtained for 160 concepts. A basic science expert (R.C., a physiologist) and a clinical expert (D.S., a general internist) independently selected from this list concepts each considered to have the greatest ramifications for facilitating the learning of other concepts in the basic sciences or for promoting a good understanding of underlying processes important in clinical medicine. Their lists contained 17 concepts in common; the concepts on the two lists showed significant agreement beyond chance ($p < .05$, McNemar test). These 17 concepts, considered important by both a basic scientist and a clinical expert, were included on the questionnaire. Ten additional concepts were chosen, five from the clinician's list and five important but less difficult ones from the original list of 160. The latter were included to provide a balance of concept difficulty within the questionnaire. The 27 concepts were included on the questionnaire in randomized order.

Designing the Questionnaire

The questionnaire consisted of four major sections plus questions about the respondent's discipline and teaching background. The first section contained the 27 biomedical concepts and four questions about each: (1) was the faculty member involved in a substantial way in teaching the concept; (2) a rating of how difficult the concept is for medical students to learn and understand during basic science years (on a scale from 5 = very difficult to 1 = not very difficult); (3) a rating of how difficult the concept is for students to recognize and/or apply in clinical situations (identical five-point scale); and (4) how important the concept is to the practice of medicine (scale from 5 = very important to 1 = not very important). The option "Cannot estimate" was included for the three rating questions.

The second and third sections of the questionnaire asked faculty to choose concepts that should be the subject of research on students' learning and understanding of medical concepts. In section two, these were chosen from the 27 concepts provided; section three provided an opportunity for faculty to nominate additional concepts. The fourth section asked for more detailed discussion of any two concepts selected by the faculty member. The questionnaire was pilot-tested, which led to development of an eight-page questionnaire format that was estimated to require approximately 30 minutes to complete.

Sampling

A two-stage sampling procedure was used to identify appropriate faculty respondents. In the first stage, the deans of all U.S. and Canadian medical schools were sent a letter in 1987 explaining the purpose of the project and asking them to nominate six faculty at their institutions with interest and experience in the education of medical students. (The authors' medical school was not included because several of the faculty had participated in research activities, including generating the concepts.) The entire range of basic and clinical science disciplines was systematically sampled across schools such that, within each school, the dean was asked to provide names for three basic and three clinical science disciplines. In the second stage, the faculty members suggested by the deans were sent a cover letter and the questionnaire. Follow-up efforts consisted of a postcard reminder approximately one month after receipt of the first mailing and a second complete mailing approximately two months later. Recipients of the second questionnaire were asked to return the cover letter sent with the questionnaire if they did not plan to respond.

Analysis

Basic descriptive statistics included proportions, means, and medians. Statistical analyses were performed using analysis of variance followed by Tukey comparisons (HSD), chi-square and Wilcoxon rank sum tests, and nonparametric correlations; only those results statistically significant at $p < .05$ are reported.

The 247 new concepts nominated by respondents in part three of the questionnaire were categorized by one of the authors (RC). All of these concepts were reanalyzed three months later by the same author without reference to the original analysis, in order to establish the consistency of the categorization. The same major categories established in the first analysis were independently established during the second analysis. Furthermore, there was an 80% agreement between these two separate categorizations of the new concepts nominated by the survey respondents.

Results

One-hundred-fifteen of the deans responded (representing 82% of the surveyed medical schools), and provided names of 324 faculty members from seven different basic science disciplines and 318 faculty from seven different clinical specialties, a total of 642. Of these, 427 (67%) responded. In this group, 44 stated they were not qualified to complete the questionnaire, and 37 opted not to fill out and return the second mailing; the remaining 346 questionnaires (54% of the original sample) were analyzed. Responses from one or more faculty were received from each medical school whose dean had nominated faculty to respond to the survey. The response rates from faculty in the basic science disciplines ranged from 23% in the behavioral sciences to 73% in physiology; in clinical specialties, the range was from 44% in psychiatry to 74% in internal medicine. The overall response rates did not differ significantly between the basic science faculty (56%) and the clinical faculty (52%).

The mean numbers of years the respondents had taught in a medical school were 17.7 years for the basic science faculty (SD = 8.0 years) and 12.9 years for the clinical faculty (SD = 8.0 years). In addition to extensive educational experience, 88% of these faculty reported that at the time of the survey or in the past they had had some level of administrative responsibility for curriculum.

Table 1 lists the 27 concepts included on the questionnaire, ranked by the relative magnitudes of the mean ratings of importance they were perceived to have "to the practice of medicine." Also given are the rankings of the mean ratings assigned to each concept with respect to both difficulty in the basic science years and difficulty of clinical application. Finally, the table presents the rankings (by the number of times nominated) of the concepts recommended as important as subjects of further research about students' learning. The mean ratings upon which the first two sets of rankings (importance and difficulty) were based are also presented. All 27 concepts were considered to be important in the practice of medicine, based upon the rating scale; the mean ratings varied from 4.6 for acid–base, electrolyte balance to 2.7 for vascular steal phenomena (5 = very important, 1 = not very important). Standard errors on all rating scales were approximately 0.07; thus, 99% confidence intervals for the mean ratings are plus or minus approximately 0.18.

The respondents were categorized according to several dimensions for statistical comparisons: whether their disciplines were basic science or clinical; whether their graduate degrees were Ph.D. (or Sc.D.), M.D., or Ph.D./M.D.; and by their disciplines or specialties. Because of the redundancy in these comparisons and the general consistency in the responses, only the comparisons for basic science versus clinical science are reported in Table 1. There were few differences between the basic science and clinical faculty members' perceptions of the importance of the concepts. They agreed upon the importance of all concepts except three: mechanisms by which neurotransmitters and other hormones affect emotion, biochemical mechanisms of turning food into energy, and the complement system. In all three cases the basic science faculty rated these concepts as more important than did the clinical faculty.

The mean ratings of difficulty in learning the concepts in the basic sciences varied from the most difficult, renal concentrating mechanisms (3.9), to the least difficult, anatomy of the spinal cord in relationship to dermatomes and vascular steal phenomena (2.2) (5 = very difficult, 1 = not very difficult; Table 1). There were 20 concepts rated similarly by basic science and clinical faculty, and seven concepts for which the mean ratings of difficulty differed between them. The basic science faculty rated ventilation/perfusion relationships as more difficult to learn in the basic sciences than did the clinical faculty. The clinical faculty, on the other hand, rated mechanisms by which hypertrophies develop; factors contributing to opposition to blood flow; complement system; hormonal aspects of digestion, absorption, and gastrointestinal motility; clonal nature of the immune response; and biochemical mechanisms of turning food into energy all as more difficult to learn in the basic sciences than did basic science faculty.

The mean ratings of difficulty in recognizing and/or applying the concept in clinical situations ranged from 3.7 for the most difficult, acid–base, electrolyte balance, to 2.3 for the least difficult, vascular steal phenomena. Again, there was general agreement on the difficulty levels of concepts, with disagreement on only five. In each instance the clinical faculty rated the concept to be more difficult to apply in clinical situations than did the basic science faculty. The five concepts were spinal cord anatomy in relationship to dermatomes; relationship of reticular activating system and state of arousal to other higher neurologic functions; complement system; hormonal aspects of digestion, absorption, and gastrointestinal motility; and mechanisms of muscle activation and control of contraction.

When distinctions between basic

Table 1

Rankings and Mean Ratings of 27 Biomedical Concepts Rated by Selected Science and Clinical Faculty According to "Importance to Practice of Medicine," Difficulty for Students to Learn and/or Apply, and Importance as Subjects of Further Research; 115 North American Medical Schools, 1987*

	Ranking by Magnitude of Mean Rating						
	Importance		Basic Science Difficulty		Clinical Difficulty		Ranking by No. of Times Concept Recommended for Research‡
Concept	Ranking	Mean Score†	Ranking	Mean Score†	Ranking	Mean Score†	
Acid-base, electrolyte balance	1.0	4.6	2.0	3.8	1.0	3.7	1.0
Mechanisms of hypoxemia	2.0	4.5	25.0	2.4	26.0	2.4	6.0
Feedback systems in endocrinology	3.0	4.2	24.0	2.5	24.5	2.6	4.0
Central control of systemic blood pressure	5.0	4.1	18.5	2.9	18.5	2.9	7.0
Interlocking regulation/control of cardiac output and central venous pressures	5.0	4.1	11.5	3.3	11.0	3.3	9.0
Alpha and beta receptors: distribution, mechanisms of action and interaction in various tissues	5.0	4.1	11.5	3.3	11.0	3.3	3.0
Serum sodium osmolarity and controls	7.5	4.0	8.0	3.4	7.0	3.4	
Ventilation/perfusion relationships and mismatches	7.5	4.0	8.0	3.4§	7.0	3.4	
Glomerular filtration rate in relation to serum creatinine and creatinine clearance	9.0	3.9	21.0	2.7	21.5	2.7	
Statistical properties of diagnostic tests and their use in revising estimates of the presence of disease	10.5	3.8	13.5	3.2	13.0	3.2	2.0
Nutrient, electrolyte, and water exchange across the cell membrane	10.5	3.8	8.0	3.4	7.0	3.4	12.5
Factors contributing to opposition to blood flow; e.g., resistance, compliance, inertance	12.5	3.6	13.5	3.2‖	14.0	3.1	
Renal concentrating mechanisms and the counter-current exchange system	12.5	3.6	1.0	3.9	2.0	3.6	10.5
Hormonal aspects of digestion, absorption, and gastrointestinal motility	15.5	3.5	18.5	2.9‖	16.0	3.0‖	
Oxyhemoglobin dissociation curve and factors that affect it	15.5	3.5	16.5	3.0	16.0	3.0	
Control and coordination of body movement: mechanisms and dysfunctions	15.5	3.5	8.0	3.4	11.0	3.3	
Anatomy of the spinal cord in relationship to dermatomes	15.5	3.5	26.5	2.2	21.5	2.7‖	
Production and function of various types of prostaglandins	19.0	3.4	21.0	2.7	21.5	2.7	12.5
Lung surfactant and reduction of surface tension in alveoli	19.0	3.4	21.0	2.7	21.5	2.7	
Complement system: classic pathway, alternate pathway	19.0	3.4§	3.0	3.6‖	3.5	3.5‖	
Clonal nature of the immune response: one cell-one specificity	21.0	3.3	16.5	3.0‖	16.0	3.0	10.5
Mechanisms by which neurotransmitters and other hormones affect emotion	22.0	3.2§	4.5	3.5	7.0	3.4	5.0
Biochemical mechanisms of turning food into energy that is useful at the cellular level, e.g., Krebs cycle, lipolysis, maintenance of ATP	23.0	3.0§	8.0	3.4‖	7.0	3.4	8.0
Relationship of reticular activating system and state of arousal to other higher neurologic functions	24.5	2.9	4.5	3.5	3.5	3.5‖	
Mechanisms by which hypertrophies develop in different organs	24.5	2.9	23.0	2.6‖	24.5	2.6	
Mechanisms of muscle activation and control of contraction	26.0	2.8	15.0	3.1	18.5	2.9‖	
Vascular steal phenomena	27.0	2.7	26.5	2.2	27.0	2.3	

*The authors, through interviews with medical school teachers and researchers, identified 160 concepts "important to the practice of medicine" and agreed upon 27 important ones as the basis of a questionnaire sent to selected faculty at all North American medical schools. The faculty were asked to rate each concept according to its importance to the practice of medicine and its difficulty (in both basic science and clinical learning). They also were asked to choose concepts that should be the subject of future research on students' learning and understanding of those concepts. A total of 346 of the questionnaires were completed (54% of the original sample).

†Range of scores was from 1 = not very to 5 = very.

‡In this column, the rankings are according to the number of times each concept was proposed as a candidate for further research. All concepts nominated by 25 or more faculty are included.

§Concepts basic scientists rated higher.

‖Concepts clinical scientists rated higher.

science and clinical faculty were ignored, there was strong agreement that concepts difficult to learn in the basic sciences are also difficult to recognize and apply in clinical situations. Except for spinal cord anatomy, the differences between mean ratings on difficulty in clinical situations and in basic science were 0.3 points or less on the rating scale. This interpretation was strengthened by the finding that the rank order correlation between ratings of difficulty in basic science and difficulty in clinical application was .87.

Comparisons were also made between the faculty who indicated they were involved in teaching each concept and those who did not teach the concept. There was a tendency for those who taught the concept to rate it higher in importance than did those who were not teaching it; this difference was observed for 17 of the 27 concepts. However, there were few differences between teachers and nonteachers in their ratings of difficulty. On two concepts the teachers rated the concept more difficult, while on three others they rated it less difficult; there were no significant differences in the remaining 59 comparisons.

In order to summarize the respondents' ratings for the relative importances of the concepts and their difficulties for students to learn (in basic sciences and to apply in clinical situations), the median ratings were considered the dividing points between high and low on these rating scales. For importance, this point was 3.5; for difficulty, it was 3.0.

Based on this classification, nine concepts were judged to have the combined qualities of being rated high in importance to the practice of medicine, high in difficulty to learn in the basic sciences, and high in difficulty to apply in clinical situations. These concepts were acid-base, electrolyte balance; regulation of cardiac output; alpha and beta receptors; statistical properties of diagnostic tests; nutrient, electrolyte exchange; renal concentrating systems; serum sodium osmolarity and controls; factors related to opposition of blood flow; and ventilation/perfusion relationships. The first six concepts on this list were among the 13 concepts (also in the table) that were selected by 25 or more faculty for further research on how students learn and understand them. The other concepts recommended for further research included three that were rated high in importance but relatively low in difficulty (mechanisms of hypoxemia; feedback systems in endocrinology; and central control of systemic blood pressure); two that were rated relatively low in importance but high in difficulty to learn and apply (mechanisms by which neurotransmitters affect emotion, biochemical mechanisms of turning food into energy); and two that were judged relatively low in importance and low in difficulty (production and function of prostaglandins; the clonal nature of immune response).

The respondents nominated an additional 247 concepts (besides the 27 included on the questionnaire) that they thought were important to the practice of medicine and/or difficult to learn and apply. The footnote[*] contains a summary of the additional concepts, organized by the authors' classification scheme developed in the content analysis. (The list of all nominated concepts is available from the authors.) This classification scheme is not presented as unique or prescriptive, but is simply reported as a reasonable way of conveying the major content areas of the data, in full awareness that other basic scientists or clinicians could organize the concepts in other reasonable ways.

Eighty-nine of the nominated concepts were related to the basic sciences; of these, 48 were physiologic concepts covering a broad range of traditional topics (neurology, cardiovascular, renal, respiratory, gastrointestinal, endocrine, and others). Fourteen concepts were pharmacologic, ten were in biochemistry, a few were related to molecular or microbiological immunology, and seven were anatomic. Seventy-nine concepts were nominated on clinical science topics; 13 each were in the areas of (1) internal medicine, such as neurology — for example, mechanisms of pain — and cardiology — for example, osmotic fluid shifts; (2) clinical pharmacology, such as mechanisms of drug toxicity; (3) genetics oncology, such as genetic diagnostic probes, inborn errors of metabolism, oncogenes, and carcinogenesis; and (4) microbiology-immunology, such as vaccine design, virology, and bacteriology. Seventy-nine concepts in clinical and preventive medicine and associated topics were also nominated; of these, 23 were related to general clinical medicine and included therapies, such as antibiotic usage; procedural management, such as anticoagulant therapy and responses to IV nutrition; and addiction, such as alcohol and nicotine. Another seven were related to clinical skills, such as differential diagnosis and interviewing. Thirteen topics were in decision making, statistics, and experimental design, such as clinical decision making, incidence and prevalence, and scientific method in medicine; and another eight related to epidemiologic concepts, such as demographics, environmental exposure, and clinical epidemiology. Nineteen concepts pertaining to psychosocial issues — such as stress, compliance, and family systems — were nominated; and another nine dealt with ethics and character, such as paternalism, ethics, and physician-patient relations.

---

[*] The concepts are in the following categories; each category is followed by a number indicating how many times concepts within it were nominated: *Basic sciences* (89 nominations): physiology (consisting of neuromuscular, 8; gastrointestinal, 5; metabolism, 8; aging, 2; cardiovascular, 12; respiratory, 3; renal, 1; endocrinology, 5; and reproduction, 4); anatomy/embryology/neuroanatomy, 7; pharmacology, 14; biochemistry, 10; molecular biology, 7; and microbiology/immunology, 3. *Clinical sciences* (79 nominations): internal medicine/pathophysiology (consisting of psychological-neurology interactions, 6; and cardiology, 7); pathology, 11; pharmacology, 13; biochemistry, 3; family practice, 5; emergency medicine, 2; genetics oncology, 13; infectious disease, 6; and microbiology/immunology, 13. *Clinical and preventive medicine and associated topics* (79 nominations): clinical medicine, 23; clinical skills, 7; decision analysis/statistics/experimental design, 13; epidemiology, 8; psychosocial medicine, 19; and ethics/character, 9.

Discussion

The challenge faced by medical educators is to accomplish two seemingly conflicting goals at once. On the one hand, the knowledge overload must be trimmed, and added responsibility must be given to students for filling out their base of medical information, perhaps with the help of better technical aids. On the other hand, some fundamental building-block concepts of medicine must be understood in depth, perhaps requiring even more intensive curricular attention. There is a tension here between the less and the more.

The current survey has provided some guidance regarding biomedical concepts that might serve as foci in a curriculum. All the concepts provided in the survey were judged to be "important to the practice of medicine"; however, there were gradations in the perceived difficulties the concepts pose to students' learning and understanding. What is perhaps most striking is the degree of agreement between basic science and clinical teachers regarding both "importance" and "difficulty." There was a strong agreement between basic scientists and clinicians on the importance of concepts, and disagreements on difficulties never involved more than 25% of the concepts. There are not "two cultures" here, one basic science and the other reflecting clinical practice; this study has identified some concepts about which there is reasonable consensus. Of course, the concepts provided in the questionnaire cannot be assumed to capture the full range of concepts that medical teaching faculty consider to be especially important and difficult. The section of the survey in which respondents nominated new concepts should aid in completing this picture.

The results of the survey can provide guidance for curriculum design or reform at two complementary levels. First, the judgments of the importance of concepts provide focus for what should receive special attention in the curriculum. Second, the judgments of the difficulty of concepts indicate which concepts are going to require "special handling"—that is, concepts that are likely to be hard for students to grasp and that may easily be misunderstood.

Some precautionary notes need to be given regarding the interpretations of the results, however, and their implications for curricular design—some of which point to the need for further research. First, the level of difficulty perceived by faculty may not be the same as the actual difficulty students encounter in understanding a concept. This problem of perception points to the need for studying, in detail, students' learning and understanding of the concepts, including the kinds of mental models[14] of the concepts they acquire. This kind of work is being undertaken by the authors and their colleagues, guided by the results of the survey.[11] Second, care must be exercised in any curricular or pedagogical response aimed at addressing areas of conceptual difficulty. For example, the seemingly reasonable pedagogical approach (one that appears often to be used) of teaching a difficult idea by starting with a simplified model and "building up" can backfire; students may not progress beyond the cognitively satisfying but overly simplified model.[11,15] This problem points to the need for research on how to teach (and learn) complicated concepts, in ways that facilitate understanding without compromising the inherent complexity itself.[16,17] Finally, educators should be aware of the extent to which entrenched misconceptions of difficult concepts are likely to exist among students and how difficult these can be to change.[18] The responses to this survey, however, indicate that not all concepts viewed as important to the practice of medicine are difficult to learn and apply. Furthermore, the guidance provided by the survey in identifying potential areas of conceptual difficulty can direct research into what characteristics make a concept difficult, the kinds of misconceptions students acquire, and pedagogical approaches that can be effective.

This research was supported by a grant (#B8520001) from the Josiah Macy, Jr. Foundation and in part by a grant (#32/85A) from the National Fund for Medical Education and Citicorp. No endorsement of the content by these foundations or agencies is suggested.

References

1. Matheson, N. W., and Cooper, J. A. D. Academic Information in the Academic Health Sciences Center: Roles for the Library in Information Management. *J. Med. Educ.* **57**, Part 2 (October 1982).
2. Muller, S. (Chairman). Physicians for the Twenty-First Century: Report of the Project Panel on the General Professional Education of the Physician and College Preparation for Medicine. *J. Med. Educ.* **59**, Part 2 (November 1984).
3. Barrows, H. S., and Peters, M. J. *How To Begin Reforming the Medical Curriculum.* Springfield, Illinois: Southern Illinois University School of Medicine, 1984.
4. Friedman, C. P., and Purcell, E. F. *The New Biology and Medical Education.* New York: Josiah Macy, Jr. Foundation, 1983.
5. Norman, G. R., Muzzin, L. J., Williams, R. G., and Swanson, D. B. Simulation in Health Sciences Education. *J. Instr. Dev.* **8**(1985):11–17.
6. Hiebert, J. *Conceptual and Procedural Knowledge: The Case of Mathematics.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1986.
7. Glaser, R. Education and Thinking: The Role of Knowledge. *Am. Psych.* **39**(1984):93–104.
8. Chi, M. T. H., Feltovich, P. J., and Glaser, R. Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Sci.* **5**(1981):121–152.
9. Feltovich, P. F., Johnson, P. E., Moller, J. H., and Swanson, D. B. LCS: The Role and Development of Medical Knowledge in Diagnostic Expertise. In *Readings in Medical Artificial Intelligence: The First Decade,* W. J. Clancy and E. H. Shortlife, eds. Reading, Massachusetts: Addison-Wesley Publishing Company, 1984.
10. Clement, J. A Conceptual Model Discussed by Galileo and Used Intuitively by Physics Students. In *Mental Models,* D. Gentner and A. L. Stevens, eds. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1983.
11. Feltovich, P. F., Spiro, R. J., and Coulson, R. L. The Nature of Understanding in Biomedicine: The Deep Structure of Complex Ideas and the Development of Misconceptions. In *Cognitive Science in Medicine:*

Biomedical Modeling, D. Evans and V. Patel, eds. Cambridge, Massachusetts: MIT (Bradford) Press, 1989.

12. COD-CAS Working Group. (1986). Commentary on the Report of the Panel on the General Professional Education of the Physician and College Preparation for Medicine. *J. Med. Educ.* **61**(1986): 346–352.
13. Barrows, H. S. *How To Design a Problem-Based Curriculum for the Preclinical Years.* New York: Springer Publishing Company, 1985.
14. Gentner, D., and Stevens, A. L., eds. *Mental Models.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1983.
15. Coulson, R. L., Feltovich, P. J., and Spiro, R. J. Foundations of a Misunderstanding of the Ultrastructural Basis of Myocardial Failure: A Reciprocation Network of Oversimplications. *J. Med. Philosophy* **14** (1989):109–146.
16. Spiro, R. J. Vispoel, W., Schmitz, J., Samarapungavan, A., and Boerger, A. Knowledge Acquisition for Application: Cognitive Flexibility and Transfer in Complex Content Domains. In *Executive Control Processes*, B. C. Britton, ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1987.
17. Spiro, R. J., Feltovich, P. J., Coulson, R. L., and Anderson, D. K. Multiple Analogies for Complex Concepts: Antidotes for Analogy-Induced Misconception in Advanced Knowledge Acquisition. In *Similarity and Analogical Reasoning*, S. Vosniadou and A. Ortony, eds. Cambridge, England: Cambridge University Press, 1989.
18. Champaign, A. B., Gunstone, R. R., and Klopfer, L. E. Instructional Consequences of Students' Knowledge about Physical Phenomena. In *Cognitive Structure and Conceptual Change*, L. H. T. West and A. L. Pines, eds. Orlando, Florida: Academic Press, 1985.

Predictive Validity of the American Board of Family Practice In-Training Examination

TERRENCE M. LEIGH, Ed.D., TIMOTHY P. JOHNSON, Ph.D., and
NICHOLAS J. PISACANO, M.D.

Abstract—Most research into the validity of graduate medical education in-training examinations has focused on construct validity and concurrent validity issues. This study examined the predictive validity of the American Board of Family Practice In-Training Examination using multivariate analysis of all U.S. family practice residents who took the certification examination in either 1987 or 1988. Results of these two certification examinations were analyzed for the cohorts of physicians who had taken in-training examinations as first-, second-, and third-year residents. Multiple regression analysis showed that the composite score, and all but one part-score, of each in-training examination were independently predictive of performance on the certification examination. This study also found that the older residents did less well on the certification exam, and that men and women each did better on selected portions of the examination. The implications of these findings for in-training examinations in other specialties and for program directors and residents are discussed. *Acad. Med.* 65(1990):454–457.

For over a quarter of a century, graduate medical education has been evaluated by in-training assessment examinations administered to residents.[1] The American Academy of Orthopedic Surgery administered the first annual in-training examination in 1961, and the American College of Physicians, the most recent organization to join this process, initiated its first examination in 1988. Presently, there are 19 medical specialty boards or specialty societies that sponsor annual in-training examinations.[2]

In undergraduate medical education, the examinations of the National Board of Medical Examiners are used to assess the performances of medical students and to provide feedback to educators. In contrast, there is no common assessment mechanism for students at the graduate medical education level. The in-training assessment examination, therefore, has evolved to meet this need within most specialties. In family practice, for example, nearly every one of the 380-plus residency programs participates in the in-training assessment process. The in-training examination (ITE) becomes a mechanism to assess the cognitive knowledge of residents within residency programs, to compare performances of residents across training levels, and to measure the performances of residents against those in comparable programs in the rest of the nation. The programmatic application of the in-training examination is also important in that it provides program directors with a means to assess the relative degrees of success of the different clinical content components of the training program (for example, do the residents do as well in surgery as they do in internal medicine, and are the composite scores in these areas similar to the national averages?) For most of these examinations, there is no determination of pass or fail.

As with all examinations, however, before one can rely on the results, it must be determined that the examination itself is valid. In-training examinations are no exception; the lit-

---

Dr. Leigh is associate executive director, American Board Practice (ABFP), and Dr. Johnson is senior project coordinator, survey research laboratory, University of Illinois, Chicago. Dr. Pisacano was executive director, ABFP, from 1970 until 1990, and was secretary of the ABFP at the time of his death in March 1990.

Correspondence and requests for reprints should be addressed to Dr. Leigh, American Board of Family Practice, 2228 Young Drive, Lexington, KY 40505.

Another article about in-training examinations appears on page 475.

erature contains many studies on the issue of validation. The studies most often reported, however, have been investigations of the construct validity of these examinations. This approach assumes that the examination is valid if it is able to perform in a manner consistent with expectations. Thus, if residents' performances on the in-training examination improve with increased residency experience, the examination can be said to have construct validity. A number of medical specialties have presented evidence of the construct validity of their in-training examinations.[3–12] Several of these studies have also demonstrated that residents' examination performances in clinical subspecialty areas also improve with increased training.

Studies of the concurrent validity of in-training examinations have also been reported in the literature. An examination is considered to have this form of validity if it correlates with other measures of performance. A number of researchers have investigated the concurrent validity of in-training examination by correlating examination scores with residents' ratings by faculty.[3,10,13–16] However, results of these studies have been mixed, suggesting that the correlation between ratings and examination scores is a function of who provides the ratings and whether the ratings are a global impression of clinical ability or assessments of clinical performances in specific areas.

A third approach to evaluating in-training examinations is by assessing predictive validity. Predictive validity, simply stated, is the extent to which an examination can predict the future performances of examinees. To the authors' knowledge, there have been only three reported studies of the predictive validity of in-training examinations.[13,17,18] In each, strong evidence was found that the in-training examination was predictive of subsequent performance on the Qualifying Examination of the American Board of Surgery. However, each of these studies has examined only the bivariate analysis of data from a single residency program. The purpose of the present study was to examine the predictive validity of the American Board of Family Practice In-Training Examination based upon a multivariate analysis of national data from recent certification examinations.

Method

The data examined were scores from both the in-training examination (ITE) and the certification examination of the American Board of Family Practice (ABFP). The ITE is given to family practice residents during November in each of the three years of residency training. Residents who take the ITE remain anonymous from the Board and are identified only by a code number assigned by the residency program. For this study, each U.S. family practice residency program was sent a listing of all the residents from that program who had taken the certification exam in 1987 and 1988, in order to match the residents' certification examination scores with the scores of their three in-training examinations. The results of all three in-training examinations were assembled for those cohorts of physicians who began their residencies in 1984 and 1985. The data available for each examination included a composite score and part scores for internal medicine, surgery, obstetrics, community medicine, pediatrics, psychiatry, gerontology, and gynecology.

The ABFP certification examinations, which were administered to these cohorts in July 1987 and July 1988, served as the criterion variables. All ABFP examinations are standardized to a mean of 500 and a standard deviation of 100. For the cohort entering residency programs in 1984, complete data were available for 1,392 individuals for three in-training assessment examinations and the 1987 certification examination. For the cohort beginning residency training in 1985, comparable complete data were available for 1,420 physicians. All the residents included in this analysis were taking the certification examination for the first time.

Two demographic variables were also examined. Each physician's age (in years) and gender (0 = female, 1 = male) were included in the analysis. Previous studies had identified these variables as being significantly associated with performances on medical specialty board certifying examinations.[19] The demographic variables were included in this analysis in an effort to control for their potential influences upon the relationship between performances on the in-training and certification examinations.

Data were analyzed using several statistical methods. Initially, a one-way analysis of variance (ANOVA) was conducted to verify the construct validity of the ITE. For this set of analyses, the 1986 and 1987 ITE composite and part scores served as dependent variables. A resident's year in residency training (first, second, or third) was the independent variable in each ANOVA. Differences between mean scores on the ITE were assessed for statistical significance using Fisher's protected least-significant difference (LSD) procedure.[20]

Correlation coefficients and multiple regression analysis were subsequently employed to estimate the predictive validity of the ITE. In the regression analysis, the 1988 and 1987 certification examination composite and part scores were used as dependent variables. Scores on ITEs taken during the three years of residency training were included as independent variables, along with age and gender.

As intercorrelations among examinations will suggest, the inclusion of in-training scores from all three residency years may produce problems of multicollinearity in regression equations. A number of solutions have been suggested in the methodology literature, including the exclusion of redundant variables, combining correlated independent variables into a single indicator, and the use of ridge regression techniques.[21,22] There is, however, no single best solution to this problem. Given that multicollinearity does not bias regression coefficients but only increases standard errors,[23] ordinary least-squares regression was used in this analysis.

Table 1

U.S. Residents' Mean Scores on All Parts of the 1986 and 1987 ABFP In-Training Examination (Listed by Year of Residency), with One-Way Analysis of Variance Values[*]

	1986 In-Training Examination				1987 In-Training Examination			
	Mean Scores, by Year of Residency				Mean Scores, by Year of Residency			
	Year 1 (*n* = 2,411)	Year 2 (*n* = 2,408)	Year 3 (*n* = 2,380)	Model F-Value	Year 1 (*n* = 2,461)	Year 2 (*n* = 2,406)	Year 3 (*n* = 2,342)	Model F-Value
Internal medicine	471.2	505.7	528.8	266.6	472.1	504.6	528.6	206.4
Surgery	478.7	500.1	524.0	130.8	469.5	504.9	529.5	236.3
Obstetrics	462.9	509.7	530.3	311.4	463.1	512.8	526.5	314.1
Community medicine	480.3	501.7	521.3	104.3	493.4	500.0	511.4	20.6
Pediatrics	476.9	500.3	526.6	156.6	465.6	503.3	535.8	326.0
Psychiatry	487.1	499.5	519.0	64.6	490.7	500.0	513.0	31.0
Gerontology	477.7	502.5	523.7	135.0	489.4	501.6	515.7	42.2
Gynecology	466.1	507.2	528.8	263.5	479.3	505.3	520.6	110.0
Composite score	453.7	506.3	544.0	580.0	468.8	505.9	534.1	427.0

[*] This table allows a comparison of residents' scores nationwide on the 1986 and 1987 American Board of Family Practice In-Training Examinations. The ANOVA values indicate that the average composite score performances on the examinations, as well as the mean part-score performances, vary significantly by year of residency. (Fisher's protected least-significant difference procedure indicated that the mean scores by year of residency were all significantly different at the .05 level.) See the text for additional interpretations of these scores and for their relation to the scores of the same residents on their certification examinations, taken in either 1987 or 1988.

Results

A comparison of differences in the 1986 and 1987 ITE mean scores by residency program training level is presented in Table 1. The one-way ANOVA model F-values indicate that average composite score performances on the individual examinations, as well as mean part-score performances, varied significantly by year of residency. For all scores examined, the average performance of the residents improved as residency training and experience increased. Fisher's protected LSD procedure indicated that average scores by year of residency were all significantly different at the 0.05 level for each of the nine in-training scores examined.

For both cohorts of residents, bivariate correlations between the three separate in-training examinations and the certification examination were between 0.69 and 0.75. Intercorrelations among the three in-training examinations taken by each group of residents were found to range from 0.67 to 0.75.

Regressions of the ABFP certification examination composite score upon the scores on the three in-training examinations and demographic characteristics of residents in both cohorts were next examined. Results for each cohort indicated that each in-training examination was independently predictive of performance on the certification exam. In addition, the older residents did less well on the certification examination. The residents' gender had no effect on overall examination scores in either cohort. The two equations each accounted for approximately two-thirds of the variation in composite scores on the certifying examination.

Additional regression equations were performed in order to assess the extent to which in-training part scores in eight subject areas were predictive of corresponding part scores on certifying examinations. In all but one instance, each in-training part score included in these equations was found to be predictive of the part score in the same subject area on the certifying examination. Age had the same effect on part scores as on the certification examination composite scores. Gender, however, showed some interesting variations by subject area. For the 1987 certification examination, the women performed significantly better in the obstetrics, pediatrics, and gynecology sections of the certifying examination, while the men did significantly better in surgery and community medicine. On the 1988 examination, the men performed better on the surgery component, while the women did significantly better on the psychiatry and gynecology sections. Past research has shown gender differences in medical students' performances in specific subject areas.[24] A review of the coefficients of determination (R^2) shows that the subject-area scores from the ITE were not as strongly predictive of their corresponding certifying examination scores as were the composite scores.

Summary and Conclusions

As mentioned above, in-training examinations are useful for a variety of purposes. These examinations are used as achievement tests to obtain an indication of the extent and quality of past instruction and experience. They are also useful as ability tests to obtain a measure of current residents' knowledge and as a way to measure the performances of residents nationally. In addition, they can be used to assess the quality of training received within clinical disciplines, and they function as aptitude tests to estimate or predict a resident's future examination performance. It has been anticipated, though not demonstrated conclusively until this study, that there is a strong positive relationship between performance on ITEs and performance on certification examinations.

This national study has confirmed the construct validity of the American Board of Family Practice In-Training Examination: as the amount of family practice clinical training increases, so does performance on the ITE. For both the total examination and the subscore areas, performances are significantly different between training levels. Moreover, this study has confirmed statistically significant predictive validity. Performance on the family practice certification examination can be predicted from the ITE scores. In-training examinations taken during all three residency years were found to be independent predictors of certification examination scores.

These data serve as clear evidence that properly designed and administered in-training assessment examinations can accurately assess the cognitive knowledge of resident physicians. And, more important, defects in cognitive knowledge identified by the ITEs studied here may forecast difficulties on future certification examinations unless there is appropriate classroom and/or clinical educational intervention. Caution is advised; these data do not suggest that the in-training assessment examination should be used as the sole determinant of performance.

Obviously, the results of this study cannot be generalized to all in-training examinations. Nor would the authors concede that all in-training examinations are of the same quality or, for that matter, used for the same purposes. However, in-training assessment examinations are generally designed to gather information about a resident's cognitive knowledge within a designated specialty. At the conclusion of residency training the physician will take a certification examination in his or her specialty area. It would follow, therefore, that the in-training examination should predict with some accuracy residency graduates' chances for attaining board certification. For those specialty areas where predictive validity studies have yet to be done, the authors encourage such efforts. Much is yet to be learned about in-training assessments and the outcomes of such assessments.

While this paper presents evidence of the predictive validity of an in-training examination program, additional research is recommended. Given the necessity of maintaining the quality of graduate medical education, efforts to validate and improve one of the primary instruments of assessing it should be not only encouraged but expected. In addition to reports on the construct, concurrent, and even predictive validity of these examinations, other types of validation require investigation. Chief among these is the content validity of these examinations. While the validity of an examination can never be established with any degree of finality, ongoing efforts to investigate the degree to which these instruments are measuring what they are intended to measure can help contribute to that elusive goal.

References

1. Chaisson, G. M. The Development and Utilization of In-Training Examinations in Graduate Medical Education. *J. Med. Educ.* **53**(1978):502-504.
2. *Annual Report and Reference Handbook—1988.* Evanston, Illinois: The American Board of Medical Specialties, 1988.
3. Buckwalter, J. A., Schumacher, R., Albright, J. P., and Cooper R. R. The Validity of Orthopaedic In-Training Examination Scores. *J. Bone Joint Surg.* 63-A (1981):1001-1006.
4. Donnelly, J. E., Yankaskas, B., Gjerde, C., Wright, J. C., and Longnecker, D. P. An In-Training Assessment Examination in Family Medicine: Report of a Pilot Project. *J. Fam. Pract.* **5**(1977):987-989.
5. Donovan, A. J. An Analysis of Results of the In-Training Examination. *Am. Coll. Surg. Bull.* **70**(1985):3-6.
6. Geyman, J. P., and Brown, T. An In-Training Examination for Residents in Family Practice. *J. Fam. Pract.* **3**(1976):409-413.
7. Grosse, M. E., Cruft, G. E., and Blaisdell, F. W. The American Board of Surgery In-Training Examination. *Arch. Surg.* **115**(1980):654-657.
8. Mahour, G. H., and Hoffman, K. I. The Development and Validation of a Standardized In-Training Examination for Pediatric Surgery. *J. Pediatr. Surg.* **21**(1986):154-157.
9. Meskauskas, J. A., Newton, M., and Russell K. P. The 1973 In-Training Examination in Obstetrics and Gynecology. Obstet. *Gynecol.* **44**(1974):463-468.
10. Nattress, L. W. Orthopaedic In-Training Examination. *J. Med. Educ.* **44**(1969): 878-889.
11. Newton, M. The 1974 In-Training Examination in Obstetrics and Gynecology. *Obstet. Gynecol.* **46**(1975):356-358.
12. Rubin, M. L., The Ophthalmology Resident In-Training Examination. *Am. J. Ophthalmol.* **67**(1969):70-79.
13. Borlase, B. C., Bartle, E. J. and Moore, E. E. Does the In-Service Training Examination Correlate with Clinical Performance in Surgery? *Curr. Surg.* **42**(1984):290-292.
14. Davis, J. K., Inamdar, S., and Stone, R. K. Interrater Agreement and Predictive Validity of Faculty Ratings of Pediatric Residents. *J. Med. Educ.* **61**(1986):901-905.
15. Kastner, L., Gore, E., and Novack, A. H. Pediatrics Residents' Attitudes and Cognitive Knowledge and Faculty Ratings. *J. Pediatr.* **104**(1984):814-818.
16. Lazar, H. L., DeLand, E. C., and Tompkins, R. K. Clinical Performance Versus In-Training Examinations as Measures of Surgical Competence. *Surgery* **87**(1980): 357-362.
17. Garvin, P. J., and Kaminski, D. L. Significance of the In-Training Examination in a Surgical Residency Program. *Surgery* **96**(1984):109-112.
18. Shetler, P. L. Observations on the American Board of Surgery In-Training Examination, Board Results and Conference Attendance. *Am. J. Surg.* **144**(1982): 292-294.
19. Norcini, J. J., Fletcher, S. W., Quimby, B. B., and Shea, J. A. Performance of Women Candidates on the American Board of Internal Medicine Certifying Examination. *Ann. Intern. Med.* **102**(1985):115-118.
20. Huitema, B. E. *The Analysis of Covariance and Alternatives.* New York: John Wiley & Sons, 1980.
21. Fennessey, J., and D'Amico, R. Collinearity, Ridge Regression, and Investigator Judgement. *Sociol. Meth. Res.* **8**(1980):P 309-340.
22. Lewis-Beck, M. S. *Applied Regression: An Introduction.* Beverly Hills, California: Sage Publications, 1980.
23. Schroeder, L. D., Sjoquist, D. L., and Stephan, P. E. *Understanding Regression Analysis: An Introductory Guide.* Beverly Hills, California: Sage Publications, 1986.
24. Miller, J. M., and Smith, I. K. The Performance of Female Medical Students in an Obstetrics and Gynecology Clerkship. *J. Med. Educ.* **57**(1982):644-646.

The Impact of an AIDS Symposium on Attitudes of Providers of Pediatric Health Care

LLOYD R. FEIT, M.D., SANFORD M. MELZER, M.D.,
STEN H. VERMUND, M.D., M.Sc., and STEVEN P. SHELOV, M.D.

Abstract—Following reports of concern among health-care workers regarding the occupational risk of infection with the human immunodeficiency virus (HIV), a symposium was designed in 1987 to demonstrate to health-care providers at three hospitals in The Bronx, New York, the low risk of occupational HIV infection and techniques for avoiding infection. After the symposium, 103 of the health-care providers who had attended it completed a questionnaire assessing the impact of the symposium on their attitudes; the responses from 100 of these providers were used in this study. Twenty-nine of the responding providers reported that the symposium had increased their concerns regarding their risk of HIV infection; this group was composed of seven of the 15 medical students who responded, 12 of the 36 housestaff, seven of the 28 faculty, and three of the 21 other medical staff. The findings of the present study suggest that a symposium designed to decrease concerns of occupational HIV infection among health-care workers may have the opposite effect on some of those who attend it, especially medical students. Education alone may be inadequate to reassure some providers. The authors recommend small-group sessions addressing the emotional aspects of health-care providers' concerns. *Acad. Med.* 65(1990):461–463.

Throughout the evolution of the acquired immunodeficiency syndrome (AIDS) epidemic, there have been numerous reports of concern and fear on the part of health-care workers in contact with AIDS patients.[1-6] Among 263 houseofficers who were recently surveyed in New York, 48% of the medical and 30% of the pediatrics houseofficers reported moderate to major concerns about acquiring AIDS from their patients; 25% of all the respondents indicated that they would not continue to care for AIDS patients if given a choice.[4]

Health-care providers from three hospitals were involved in this study: the Bronx Municipal Hospital Center, the North Central Bronx Hospital, and the Montefiore Medical Center. They are all located in The Bronx, a borough of New York City with a population of 1.16 million people, and serve a population of children in which the seroprevalence for HIV antibodies in newborns is 2.3%.[7] In this setting, we have noted increasing concern among providers of pediatric health care regarding the risk of contracting HIV infection in the course of patient care. We have also encountered instances of decreased willingness on the part of these providers to care for children with AIDS. In response to these concerns, in 1987, the Division of Pediatric Education designed an educational program entitled "AIDS: A Symposium for the Health Care Professional in Pediatrics." One of the primary goals of the symposium was to decrease the level of concern among these health professionals by presenting data on the low risk of occupational HIV infection along with a discussion and demonstration of techniques to avoid occupational HIV infection. In addition, we sought to increase the willingness of providers to care for children with AIDS.

In the present report we describe the impact of the symposium as reported by the attendees on their concerns, practices, and willingness to care for children with AIDS.

Method

The symposium was widely advertised in order to attract the greatest possible number of participants. The following paragraphs present an overview of the symposium and the questionnaire we designed to evaluate it.

The AIDS Symposium

The symposium consisted of three presentations one week apart, each 90 minutes long. The series began with an update on the epidemiology and modes of transmission guidelines and techniques for preventing HIV infection in the hospital. The third session focused on the psychosocial and ethical issues encountered in treating children with AIDS. The content of the presentations varied, and they included lectures, a panel discussion, and an emotional plea by a houseofficer encouraging her colleagues to use gloves for all invasive procedures.

Evaluation

Prior to the symposium, we developed a questionnaire to assess the efficacy of the program in decreasing concerns among the attendees regarding caring for AIDS patients. We distributed the questionnaire at the beginning of the third lecture. It was designed to test the following hypotheses:

Dr. Feit is a fellow in pediatrics and cardiology, Yale–New Haven Hospital, Connecticut; Dr. Meltzer is director, Pediatrics, Swedish Hospital, Seattle, Washington; Dr. Vermund is director, Section on Epidemiology, Division on AIDS, National Institutes of Health, Bethesda, Maryland; and Dr. Shelov is director, Division of Pediatric Education, Department of Pediatrics, Bronx Municipal Hospital Center, New York City.

Correspondence and requests for reprints should be addressed to Dr. Shelov, Director, Division of Pediatric Education, Department of Pediatrics, Room 803, Bronx Municipal Hospital Center, Pelham Parkway and Eastchester Road, Bronx, NY 10461.

Other articles about the effects of AIDS on educating health-care providers appear on pages 446, 458, 464, 467, 470, 472, and 482.

Table 1

Numbers and Percentages of 100 Pediatrics Health Professionals of Different Levels of Training Whose Attitudes toward Treatment of AIDS Patients Did or Did Not Change after an Educational Symposium on AIDS, Bronx Municipal Hospital Center, 1987[*]

	Medical Students		Housestaff		Faculty		Other Medical Staff		All Respondents	
Attitude and Response	No.	%	No.	%	No.	%	No.	%	No.	%
Perceived personal risk of occupational HIV infection										
No change	8	53	22	61	20	71	13	62	63	63
Increased risk	7	47	12	33	7	25	3	14	29	29
Decreased risk	0	0	2	6	1	4	5	24	8	8
No response	0		0		0		0[†]		0[†]	
Total	15	100	36	100	28	100	21	100	100	100
Willingness to care for future AIDS patients										
No change	14	93	31	86	24	86	17	81	86	86
Increased	0	0	4	11	0	0	4	19	8	8
Decreased	1	7	1	3	4	14	0	0	6	6
No response	0		0		0		0[†]		0[†]	
Total	15	100	36	100	28	100	21	100	100	100
Intended to take precautions in future										
No change	3	20	10	28	15	54	11	52	39	39
Increased precautions	12	80	26	72	13	46	10	48	61	61
Decreased precautions	0	0	0	0	0	0	0	0	0	0
No response	0		0		0		0[†]		0[†]	
Total	15	100	36	100	28	100	21	100	100	100

[*] A symposium was given to pediatric health-care providers, in part to present data demonstrating the low risk of occupational HIV infection and techniques for avoiding infection. The results presented in this table suggest that such a symposium, presented to decrease concerns, may have the opposite effect on some of those who attend it.

[†] One person in this "other medical staff" group did not respond, but for other reasons the other data provided by that person had to be excluded from the study results.

1. Attendance at the symposium would decrease the attendees' concerns regarding their personal risks of becoming infected with human immunodeficiency virus (HIV) as a result of patient care.
2. Attendees would report that they intended to increase the routine precautions they would take to avoid HIV infection during the course of patient care.
3. Attendance at the symposium would increase the participants' willingness to care for AIDS patients in the future.

Statistical Analysis

The questionnaires were scored on a three-point Likert scale. Chi-square analysis with the Yates' continuity correction as well as two-tailed Fisher's exact test (FET) and McNemar's binomial test was used for 2 × 2 contingency tables as appropriate.

Results

Given our desire to evaluate the impact of this symposium on health-care providers with different levels of training, it was gratifying to see the diversity among those who attended, as described in the following paragraphs.

Respondents

A total of 103 health-care providers completed the questionnaire, of whom 88 attended two or three sessions in the series, while 15 attended one session only. Data from three of the respondents were not usable for this study.

As shown in Table 1, the respondents, categorized by level of training, included 15 medical students, 36 housestaff (including fellows and chief residents), 28 faculty, and 21 other medical staff (nurses, social workers, and administrative personnel). Sixty-nine were women and 50 were between 21 and 30 years of age; the remaining respondents were older.

Respondents' AIDS Experience

Fifty eight of the respondents reported having cared for ten or fewer patients with AIDS over the previous five years, while 45 had cared for more than ten AIDS patients in that period. Twenty-four respondents reported having had percutaneous exposure to the blood or secretions of an AIDS patient, while 75 denied having such exposure (four subjects did not respond). Of the 24 respondents reporting percutaneous exposure there were two (8%) medical students, 12 (50%) housestaff, nine (38%) faculty, and one (4%) other medical staff members. By level of training, 35% of the housestaff, 32% of the faculty, 13% of the medical students, and 5% of the other medical staff reported percutaneous exposure to the blood or secretions of an AIDS patient.

The Symposium's Impact

The respondents were asked, "as a result of this symposium, have your concerns changed regarding your personal risk of becoming infected with HIV as a result of patient care activities?" Twenty-nine (29%) of the respondents reported increased concern about their personal risks of HIV infection. Increased concern was related inversely to level of training. Decreased concern was noted by less than 4% of the medical student, faculty, and housestaff groups combined. Within the other medical staff group, 24% reported decreases in their concerns ($p < .01$ by FET).

Sixty-one of the respondents (including 80% of the medical students and 72% of the housestaff) reported that as a result of the symposium they planned to increase the precautions they would take to avoid HIV infection when providing patient care in the future. The remaining 39 respondents reported that the symposium had no effect on the precautions they would take during future patient care. None stated they would decrease their precautions. Persons with increased concerns with regard to occupationally acquired HIV infection were far more likely to report intentions to increase precautions in the future ($p < .001$ by McNemar test).

Eighty-six of the respondents reported that as a result of the symposium, their willingness to care for AIDS patients in the future had not changed. Six respondents (four faculty, one houseofficer, and one medical student) felt the symposium had made them less willing to care for future AIDS patients. Of these six, none had a history of percutaneous exposure to the secretions of an AIDS patient, and only one had cared for more than ten patients with AIDS. All six reported both increased anxiety and increased concern about occupational HIV risk as a result of attending the symposium.

Discussion

The present study illustrates that an educational program designed to decrease concerns regarding occupational acquisition of HIV infection and increase willingness to care for AIDS patients may have the opposite effect on the attendees. Although the symposium was designed with a goal of alleviating concerns regarding occupational HIV infection, it served to increase concerns in almost 30% of the attendees.

With the not-unexpected finding of increased risk perception among students and houseofficers, another goal of the symposium was to encourage participants to increase the precautions they would take to avoid HIV infection while caring for all patients. Over 60% of the respondents reported that as a result of the symposium they intended to increase precautions during patient care. Medical students, the group with the greatest concern about HIV infection, were the most likely to report an intent to increase precautions as a result of the symposium.

Despite an increase in general anxiety and concerns about personal risk, 86% of the respondents reported that their willingness to care for AIDS patients in the future was unchanged. Nevertheless, these findings suggest that dissemination of information alone may be inadequate in addressing the concerns, both warranted and unwarranted, among health-care workers who care for patients with AIDS.[8] More data are needed that define the specific aspects of an AIDS educational curriculum that are most effective in changing attitudes and behaviors, reducing anxiety, and disseminating technical information. Approaches to AIDS education that address emotional aspects of health-care problems and provide support for health-care providers — with specific targeting of medical students — should be explored.[8] A prospective evaluation of the effect of such a symposium on actual rather than intended behaviors may be valuable in planning and evaluating future programs.

References

1. Zuger, A., and Miles, S. Physicians, AIDS, and Occupational Risk. Historic Traditions and Ethical Obligations. *JAMA* **258**(1987):1924 - 1928
2. Wachter, A. U. The Impact of the Acquired Immunodeficiency Syndrome on Medical Residency Training. *N. Engl. J. Med.* **314**(1986):177 - 180.
3. Douglas, C. J., Kalman, C. M., and Kalman, T. P. Homophobia among Physicians and Nurses: An Empirical Study. *Hosp. Community Psychiatry* **36**(1985):1309 - 1310.
4. Link, R. N., Feingold, A. R., Charap, M. H., and Shelov, S. P. Concerns of Medical and Pediatric House Officers about Acquiring AIDS from Their Patients. *Am. J. Public Health* **78**(1988):455 - 459.
5. Blumfield, M., et al. Survey of Attitudes of Nurses Working with AIDS Patients. *Gen. Hosp. Psychiatry* **9**(1987):58 - 63.
6. Kelly, J. A., St. Lawrence, J. S., Smith, S., Hood, H. V., and Cook, D. J. Medical Students' Attitudes Toward AIDS and Homosexual Patients. *J. Med. Educ.* **62**(1987):549 - 556.
7. Novick, L. F., Berns, D., Stricof, R., et al. HIV Seroprevalence in Newborn Infants in New York State. IV. International Conference on AIDS, Stockholm, Sweden. Abstract #7221, Book 1, p 433, 1988.
8. Stevens, L. A., and Muskin, P. R. Techniques of Reversing the Failure of Empathy toward AIDS Patients. *J. Am. Acad. Psychoanal.* **15**(1987):539 - 551.

Attitudes of Medical School Faculty and Students toward Acquired Immunodeficiency Syndrome

THEODORE B. FELDMANN, M.D., ROGER A. BELL, Ed.D.,
JUDITH J. STEPHENSON, S.M., and FRANCES E. PURIFOY, Ph.D.

Abstract—Attitudes of faculty and students will influence the success of educational programs to address acquired immunodeficiency syndrome (AIDS). In 1988–89, 170 medical school faculty and 227 students completed needs assessment questionnaires at the University of Louisville, and the results were compared to assess the respondents' attitudes about AIDS. Both the faculty and the students were found to be fearful of infection, had strong negative attitudes toward AIDS patients, and had low levels of confidence in the current knowledge about AIDS. Both cognitive and affective training strategies are recommended to help learners distinguish unwarranted from warranted fears of infection and to increase their empathy for patients with AIDS. *Acad. Med.* 65(1990):464–466.

Acquired immunodeficiency syndrome (AIDS) has reached epidemic proportions, with over 450,000 cases projected to exist in the United States by 1993.[1] This magnitude of cases makes it imperative that physicians be well trained regarding this illness; to accomplish this presents a major challenge to medical educators.

Several studies have reported negative attitudes toward AIDS patients and homosexual patients among physicians, housestaff, registered nurses, and medical students.[2-6] Fear of contagion and homophobia greatly affected the attitudes reported in these studies. To further examine these findings, the AIDS Health Education Program of the University of Louisville's Department of Psychiatry and Behavioral Sciences, which provides AIDS training to health-care providers, surveyed all medical school faculty and students regarding their attitudes about AIDS. To our knowledge, this is the only such survey at a medical school in the United States.

Method

A needs assessment was sent to 248 faculty and 485 students at the University of Louisville School of Medicine during the 1988–89 academic year. A total of 170 faculty and 227 students returned the questionnaire (response rates of 69% and 47%, respectively). The current report compares the responses made by these faculty and students to 20 attitude items related to three primary areas: (1) HIV transmission and fear of contagion; (2) feelings about AIDS and homosexuality; and (3) legal and ethics issues related to the care of patients with AIDS.

Responses were scaled as follows: 0 = no opinion; 1 = strongly disagree; 2 = disagree; 3 = agree; and 4 = strongly agree. In order to compare the responses of the faculty and students to these 20 items, chi-square analyses were performed using three categories: "No opinion;" "yes" (created by adding all "agree" and "strongly agree" responses); and "no" (created by adding "disagree" and "strongly disagree" responses). Acceptance of significant differences was based on an alpha level of .05. Trends toward significance, however, are noted if $p < .10$.

Results and Discussion

The mean age of the faculty members in our sample was 45.1 years and that of the students was 26.3 years. More of the faculty were men and married than were the students.

The responses to the 20 attitude items, presented in Table 1, indicate that at the time of our study large numbers of the faculty and students surveyed (1) were fearful of infection; (2) felt uncomfortable dealing with AIDS patients; (3) had negative attitudes and biases toward those with AIDS; and (4) experienced confusion and discomfort when confronted with the legal and ethics issues associated with AIDS. Very little difference in attitudes was found between the two groups. The responses to only two items showed significant attitude differences ($p \leq .05$), and the differences shown by the responses to another four items approached significance. In reviewing these six items, it appears that the students were more fearful of infection and felt more strongly that physicians should have the right to refuse treatment to persons with AIDS. On the other hand, the students were more sensitive to AIDS patients' right to die and to the problems of widespread HIV testing.

A number of factors may account for the similarities in responses of the two groups. Since criteria for medical school selection are determined by the faculty, the students selected are likely to share many attitudes held by the faculty. Faculty also serve as role models, influencing students' views of clinical, social, and ethics issues. Finally, the attitudes of both students and faculty tend to mirror those of

---

Dr. Feldmann is co-director, AIDS Health Education Program (AHEP), and assistant professor of psychiatry; Dr. Bell is director, AHEP, and professor of psychiatry; Ms. Stephenson is a biostatistician; and Dr. Purifoy is assistant director, AHEP; all with the Department of Psychiatry and Behavioral Sciences, University of Louisville School of Medicine, Louisville, Kentucky.

Correspondence and requests for reprints should be addressed to Dr. Feldmann, Department of Psychiatry and Behavioral Sciences, University of Louisville School of Medicine, Louisville, KY 40292.

Other articles about the effects of AIDS on educating health-care providers appear on pages 446, 458, 461, 467, 470, 472, and 482.

Table 1

Percentages of Faculty and Students Who Said "Yes," "No," or Had No Opinion in Response to 20 Items Concerning Their Attitudes toward Aspects of AIDS; University of Louisville School of Medicine, 1988–89[*]

	Faculty				Students				
Item	*n*	% Yes	% No	% No Opinion	*n*	% Yes	% No	% No Opinion	*p*
Have knowledge to protect self from AIDS	170	87.6	11.2	1.2	226	91.2	8.4	0.4	[‡]
Would prefer to avoid caring for PWA[†]	170	50.6	40.0	9.4	225	43.6	48.4	8.0	[‡]
Health providers should be able to refuse care for PWA[†]	170	33.5	59.4	7.1	217	45.1	48.2	6.6	<.07
Feel am at risk getting AIDS because of job	170	16.5	80.6	2.9	226	26.1	71.7	2.2	<.08
Sought opportunity to work with PWA[†]	170	7.6	66.5	25.9	226	13.7	69.9	16.4	<.03
Feel comfortable conversing with PWA[†]	170	88.8	10.0	1.2	226	87.2	9.3	3.5	[‡]
Feel comfortable shaking hands with PWA[†]	170	83.5	14.7	1.8	225	86.2	12.0	1.8	[‡]
Feel comfortable sharing house with PWA[†]	169	33.7	56.8	9.5	226	33.2	60.6	6.2	[‡]
Feel more negative about homosexuality since AIDS	169	29.6	63.3	7.1	226	34.5	58.8	6.6	[‡]
Believe working with PWA[†] could give AIDS to family	169	23.7	69.2	7.1	226	25.2	68.6	6.2	[‡]
Believe could get AIDS while providing routine care to PWA[†]	170	25.3	65.3	9.4	226	39.4	58.0	2.7	<.001
HIV testing should be done only with informed consent	170	42.9	50.0	7.1	225	32.4	61.8	5.8	<.07
Terminally ill PWA[†] have right to electively terminate lives	169	58.6	29.6	11.8	226	65.9	24.3	9.7	[‡]
Would assist terminally ill PWA[†] in electively ending live	169	16.0	72.8	11.2	226	25.2	63.7	11.1	<.09
Physician should give HIV test to patient if thinks patient engaging in unsafe sex	168	45.2	47.0	7.7	226	43.8	49.6	6.6	[‡]
Concerned that in future will find AIDS transmitted in ways now thought safe	170	51.8	38.8	9.4	226	56.6	34.1	9.3	[‡]
Government should distribute needles to prevent spread of AIDS	169	40.2	46.2	13.6	226	37.6	54.0	8.4	[‡]
All patients admitted to hospital should be HIV-tested	169	47.9	48.5	3.6	225	39.1	56.4	4.4	[‡]
PWA[†] engaging in unsafe sex should be incarcerated	169	63.3	28.4	8.3	226	68.6	22.6	8.8	[‡]
Rights of PWA[†] have been overemphasized	170	40.6	46.5	12.9	225	33.3	52.0	14.7	[‡]

[*] A needs assessment was sent to 248 faculty and 485 students; 170 of the faculty (69%) and 227 of the students (47%) returned it. The table presents the responses to 20 items concerning attitudes about HIV transmission and fear of contagion, feelings about AIDS and homosexuality, and legal and ethics issues related to HIV infection. As the table indicates, a few faculty and students did not respond to certain items.

[†] Patients with AIDS.

[‡] Difference between responses of students and faculty not significant.

the community at large, where fear of infection and homophobia are prominent.

Conversely, there are also several possible reasons why the students and faculty differed in their attitudes about AIDS. Fear of infection would be expected to be greater among students, who have less medical education. On the other hand, faculty in certain specialties or in the basic sciences are less likely to seek opportunities to work with AIDS patients by virtue of their chosen fields. The attitudes of clinical versus preclinical faculty about AIDS merit further study. Finally, by virtue of age and generational differences, students are likely to view social and ethics issues differently from the way their older faculty counterparts view them.

While there was a general feeling that the groups had enough knowledge to protect themselves from infection, there remained concern that working with AIDS patients carried risk. These concerns probably influenced many to avoid working with persons with AIDS, and to feel that they should be able to refuse to treat AIDS patients. This finding is important because reluctance to work with AIDS patients has the potential for negative impact on the quality of care provided.

Although the majority of both faculty and students said they felt comfortable conversing with or shaking hands with a patient with AIDS, a minority said they would feel comfortable living in the same house with someone with AIDS. This difference may reflect an association between degree of discomfort and closeness of contact and therefore an underlying fear of infection. To what extent homophobia or other stigmas such as illegal drug use, which are commonly associated with persons with AIDS, influenced the responses to this item is not known and obviously warrants further study. The concern that future modes of HIV transmission will be found also reflects an underlying fear of infection and implies a low level of confidence in the current knowledge about AIDS. Other studies have also shown that fear of infection is an overriding concern among physicians and medical students.[4,7]

That one-third of those surveyed felt more negatively about homosexuality as a result of AIDS supports findings that have been reported elsewhere.[2-5] These feelings will present significant problems for physicians treating patients who became infected as a result of homosexual behavior. In addition, the fact that over one-third of those surveyed felt that the rights of AIDS patients had been overemphasized represents a potential source of resistance to training efforts and is disturbing because more faculty than students felt this to be true. Faculty may thus feel ambivalent about including AIDS in the curriculum. Although the confidentiality of the survey results was guaranteed, negative feelings about AIDS may also have resulted in our relatively low return rates, particularly among the students.

Our findings have several implications for curriculum design and development. First, AIDS training efforts need to address both the warranted and the unwarranted fears that medical students and faculty may have about working with patients with AIDS. Lectures should include information about the actual risks involved in patient care activities that involve exposure to infected blood and ways that such exposure can be prevented. However, training should include affective as well as cognitive training strategies to help learners process the information being presented and to help alleviate unwarranted fears of HIV transmission from casual contact. We have found small-group discussions, which allow health-care providers to share their doubts and fears with colleagues, to be effective in this regard.

Likewise, increasing comfort levels through interactions with AIDS patients is very important. Personalizing this illness for physicians and students would increase empathy for AIDS patients and offset the tendency to stigmatize those with AIDS.[8] Considerable attention also needs to be devoted to the legal and ethics issues related to the care of patients with AIDS, as well as to issues regarding homosexuality and homophobia. Finally, training efforts need to be directed at faculty as well as students, because the two groups share many of the same negative attitudes.

This project was supported by National Institute of Mental Health Contract #278-88-0015.

References

1. Curran, J. W., et al. Epidemiology of HIV Infection and AIDS in the United States. *Science* **239**(1988):610–616.
2. Kelly, J. A., St. Lawrence, J. S., Smith, S., Hood, H. V., and Cook, D. J. Medical Students' Attitudes Towards AIDS and Homosexual Patients. *J. Med. Educ.* **62**(1987):549–556.
3. Douglas, C. J., Kalman, C. M., and Kalman, T. P. Homophobia among Physicians and Nurses: An Empirical Study. *Hosp. and Community Psychiatry* **36**(1985):1309–1311.
4. Merrill, J. M., Laux, L., and Thornby, J. I. AIDS and Student Attitudes. *South. Med. J.* **82**(1989):426–432.
5. Royse, D., and Birge, B. Homophobia and Attitudes Towards AIDS Patients among Medical, Nursing, and Paramedical Students. *Psychological Rep.* **61**(1987):867–870.
6. Barkley, S. E., Mason, J. L., and Kappleman, M. M. An Assessment of Entering Freshman Medical Students' Knowledge of Attitudes Toward AIDS. In *Research in Medical Education: 1988. Proceedings of the Twenty-Seventh Annual Conference*, Debra S. Dabney, compiler, pp. 221–225. Washington, D.C.: Association of American Medical Colleges, 1988.
7. Richardson, J. L., Lochner, T., McGuigan, K., and Levine, A. M. Physician Attitudes and Experience Regarding Care of Patients with Acquired Immune Deficiency Syndrome (AIDS) and Related Disorders (ARC). *Med. Care* **25**(1987):675–685.
8. Stevens, L. A., and Muskin, P. R. Techniques for Reversing the Failure of Empathy Towards AIDS Patients. *J. Am. Acad. Psychoanal.* **15**(1987):539–551.

Medical and Nursing Students' Attitudes about AIDS Issues

BONNIE J. TESCH, M.D., DEBORAH E. SIMPSON, Ph.D., and
BARBARA D. KIRBY, M.D.

Abstract—A 12-item questionnaire was administered in late 1987–early 1988 to 445 medical students, 133 medical school applicants, and 111 nursing students to assess any differences in their attitudes toward medicine-related AIDS issues. These groups were also given a 31-item test of their knowledge of AIDS issues. Significant differences by levels of knowledge were obtained for eight of the 12 attitude items. For example, the more knowledgeable the student, the less likely he or she was to refuse treatment to an AIDS patient, to require mandatory AIDS testing of physicians, or to require medical personnel to wear gloves. The findings strongly suggest that education has an important role in changing attitudes about AIDS in a direction that fosters better health care for AIDS patients. *Acad. Med.* 65(1990):467–469.

Today's students in the health professions can expect to encounter AIDS patients with ever-increasing frequency even in the early years of their professional careers. The attitudes that these future health-care providers develop towards the disease AIDS and towards individuals with AIDS will ultimately affect their ability to provide competent and compassionate care to AIDS patients.

Recently, a number of surveys and questionnaires have been developed to gather attitudinal information about AIDS from a broad range of subject groups including physicians,[1–3] health-care providers,[4] and college students.[5–7] The domains covered by these investigators include the availability of information about AIDS and information sources,[6] misconceptions about casual transmission of AIDS,[4–7] fear of AIDS,[5] and attitudes toward homosexuality.[2,3,5,8] The authors have found few articles that compare the attitudes of health-professions students toward AIDS-related medical issues. The present article reports the results of a survey designed to assess current and future medical students' and current nursing students' attitudes toward the medical issues related to AIDS. The authors chose this topic because the competent and compassionate care of AIDS patients requires that two of the key members of the health-care team—the nurse and the physician—understand while still in training the conflict arising from possible differences in their attitudes towards AIDS health-care issues.

Method

A 12-item questionnaire was developed, drawing from items included in an opinion poll of internal medicine specialists.[9] The students were asked to indicate whether they agreed or disagreed with each of the 12 statements, which were designed to elicit their attitudes about AIDS issues and were part of a larger instrument to assess students' knowledge about these issues, administered between October 1987 and January 1988. Members of three major student health-care provider groups were surveyed: (1) 133 applicants to medical school at the time of their on-campus interviews at the Medical College of Wisconsin; (2) 445 first-, second-, and third-year medical students at that school; and 111 second-year (preclinical) nursing students at the University of Wisconsin–Milwaukee, School of Nursing. (Hereafter in this article, all these types of providers are referred to as "the students" when the entire group is being indicated.) All the students in each group voluntarily completed the questionnaire, during regularly scheduled class time, at registration, or during free time on medical school interview days. The responses were given anonymously and coded only by cohort status.

Three analyses were performed. First, response frequencies were calculated. Second, chi-square tests of association were conducted to determine whether there were significant differences between the responses of the groups. Third, 3 × 3′ chi-square tests of association were conducted to determine whether there were any significant relationships between the students' knowledge about AIDS and their opinions about it. The students' knowledge about AIDS was ascertained by their performances on a 31-item test that assessed five AIDS-related knowledge domains: epidemiology, risk factors, signs and symptoms of clinical illness, mechanisms of transmission, and testing and interpretation. The responding students were classified into one of three groups based on their knowledge of AIDS: those who answered 74% or more of the items correctly were placed in the "high" group ($n = 216$); those who answered between 61% and 71% of the items correctly were placed in the middle group ($n = 325$); and those who answered 58% or less of the items correctly were placed in the low group ($n = 175$).

Results

For each item in the attitudes questionnaire, the proportion of the students who agreed with the statement (by marking "yes") was calculated. The students' agreement on re-

---

Dr. Tesch is assistant professor medicine and director, Student Health Services; Dr. Simpson is assistant professor of family medicine and director, Educational Services; and Dr. Kirby formerly was associate professor of medicine and associate dean for academic affairs; all at the Medical College of Wisconsin, Milwaukee.

Correspondence and requests for reprints should be addressed to Dr. Tesch, P.O. Box 135, Milwaukee County Medical Complex, 8700 West Wisconsin Avenue, Milwaukee, WI 53226.

Other articles about the effects of AIDS on educating health-care providers appear on pages 446, 458, 461, 464, 470, 472, and 482.

Table 1

Percentages of Medical Students, Medical School Applicants, and Nursing Students Who Answered "Yes" to 12 Questions about AIDS Issues, Medical College of Wisconsin, 1987–1988*

	Percentage Answering "Yes"			
Questions	Medical Students ($n = 445$)	Applicants ($n = 133$)	Nursing Students ($n = 111$)	All Students ($n = 689$)
Do you believe you have the right to refuse to treat AIDS patients?†	66	37	58	59
Would you refuse to treat AIDS patients?†	22	8	19	19
Should AIDS testing be mandated for some groups?	78	75	87	79
Should physicians be tested for AIDS?†	66	70	89	71
Would you restrict people with AIDS from the following occupations:				
Medical personnel?‡	53	49	66	55
Food handlers?	55	46	60	54
Teachers?	13	8	17	13
Occupations with wide-spread public contact?	23	13	11	15
Would you inform an AIDS patient's sexual partner against the patient's wishes?†	76	66	58	70
Should dental and medical personnel be required to wear gloves?	79	77	87	80
Do you believe that life and health insurance companies have the right to test applicants?	50	46	54	50
Should experimental drug testing rules be relaxed?†	81	75	53	75

*The authors administered these questions to the students and applicants in these three groups; all those who were surveyed completed the questionnaire. These responses were analyzed regarding the differences of attitudes between the groups and how the attitudes differed according to the respondents' levels of knowledge about AIDS issues, ascertained by a separate 31-item test. These responses were also compared with the responses on a similar opinion poll of internal medicine specialists.

†$p < .01$.

‡$p < .05$.

sponses was defined as either less than or equal to 20% of the students selecting the "yes" option, or greater than or equal to 80% or more of the students selecting that option. Using this criterion for agreement, the students in all three groups agreed on only three of the 12 items. Only 13% of the students felt that AIDS patients should be restricted from teaching; 15% felt that AIDS patients should be restricted from other occupations with widespread public contact. Eighty percent of the students felt that medical personnel should be required to wear gloves (see Table 1).

There was significant disagreement among the groups on six of the 12 attitude items. For example, more of the medical students felt that they had a right to—*and* would—refuse to treat an AIDS patient than did the nursing students or the medical school applicants. The medical students were also less likely than were the nursing students or the medical school applicants to feel that physicians should be tested for AIDS or that physicians with AIDS should be restricted from the practice of medicine (see Table 1).

We then grouped all the respondents according to their knowledge levels: low, middle, and high. (The resulting findings, which follow, are not found in Table 1.) Significant differences in responses according to the students' levels of knowledge were obtained on eight of the 12 attitude items. Four of these items were significant at the $p < .05$ level: the students who were more knowledgeable were less likely to refuse to treat an AIDS patient, to require mandatory testing, or to restrict AIDS patients as teachers, and would be more likely to inform a partner against the wishes of the patient. The other four attitude items achieved significance at the $p < .01$ level: the more knowledgeable students would be less likely to test physicians for AIDS, restrict AIDS patients from occupations with pubic contact, or require medical personnel to wear gloves. They would be more likely to relax restrictions on drugs for those substances that may cure or treat AIDS but that are untested.

Discussion

To provide compassionate, humane care to AIDS patients and carriers of the AIDS virus, physicians and nurses need to be knowledgeable about the disease AIDS and be aware of their attitudes toward AIDS and AIDS patients. In this survey of health-care-professions students, it was found that the more knowledgeable students had attitudes that showed a greater willingness to care for AIDS patients and less fear of these patients. The more knowledgeable students were also less supportive of mandatory testing of some groups and would not restrict AIDS patients from working in non-medicine-related occupations or with the public. The more knowledgeable students would also be less likely to require gloves and more willing to relax experimental drug rules.

It is interesting that more of the nursing students than the medical school applicants or the medical students indicated that they would test physicians for AIDS (89%, 70%, and 66%, respectively). In comparing these responses to this issue with practicing internists' responses on the same issue (obtained by the opinion poll on which the attitude questionnaire was based), this trend of lessening percentages continued, with only 39% of the practicing physicians stating that it was appropriate to test their colleagues for AIDS.[9]

In other areas, the medical students and nursing students appear to have responded more like their practicing internist colleagues, and less like the medical school applicants. For example, more of the medical students (66%) than the preclinical nursing students (58%) or medical school applicants (37%) felt they had the right to refuse to treat AIDS patients. In the internist survey, 55% of the respondents felt that they had the right to restrict treatment.[9] Twenty-two percent of the medical students, 19% of the preclinical nursing students, 15% of the internists, and 8% of the applicants said they would actually refuse treatment. The practicing internists, however, had different opinions relevant to the issue of patient confidentiality. Only 49% of the practicing internists would inform a patient's sexual partner against the patient's wishes, in contrast to 76% of the medical students, 66% of the medical school applicants, and 58% of the nursing students.

In regard to the attitude differences between the groups surveyed, it is clear that the nursing and medical students in training did have different opinions on medicine-related AIDS issues, but not in the areas regarding the right to refuse to treat or their willingness to treat AIDS patients. In contrast, the medical school applicants did differ significantly from their future colleagues on these issues.

The contrasting opinions of the medical school applicants and the finding that the more knowledgeable students had a greater willingness to care for AIDS patients, and less fear of them lends strong support to the role of education[3,10] as a method of relieving unwarranted fears of future members of the AIDS patient's health-care team.

This work was conducted through the Office of Academic Affairs, Medical College of Wisconsin. We would like to thank the staff of the University of Wisconsin-Milwaukee, School of Nursing for assistance in administering the tests to their students, especially associate dean Harriet Lazinski, her assistant, Miriam Ross, professor Flora Fernandez, and professor Mary Ann McGovern. We also thank Lesley Mack, director of admissions at the Medical College of Wisconsin, for her assistance in administering the tests to the medical students and medical school applicants.

References

1. Kelly, J. A., Lawrence, J. S., Smith, S., Hood, H. V., and Cook, D. J. Stigmatization of AIDS Patients by Physicians. *Am. J. Public Health* **77**(1987):789–791.
2. Mathews, W. C., Booth, M. W., Turner, J. D., and Kessler, L. Physicians' Attitudes toward Homosexuality—Survey of a California County Medical Society. *West J. Med.* **144**(1986):106–110.
3. Richardson, J. L., Lochner, T. McGuigan, K., and Levine, A. M. Physician Attitudes and Experience Regarding the Care of Patients with the Acquired Immune Deficiency Syndrome (AIDS) and Related Disorders (ARC). *Med. Care* **25**(1987):675–685.
4. Wertz, D. C., Sorenson, J. R., Liebling, L., Kessler, L., and Meeren, T. C. Knowledge and Attitudes of AIDS Health Care Providers before and after Education Programs. *Public Health Rep.* **102**(1987):248–254.
5. Goodwin, M. P., and Roscoe, B. AIDS: Students' Knowledge and Attitudes at a Midwestern University. *J. Am. Coll. Health* **36**(January 1988):214–222.
6. Katzman, E. M., Mulholland, M., and Sutherland, E. M. College Students and AIDS: A Preliminary Survey of Knowledge, Attitudes, and Behaviors. *J. Am. Coll. Health* **37**(November 1988):127–130.
7. Landefeld, C. S., Chren, M., Shega, J., Speroff, T., and McGuire, B. S. Students' Sexual Behavior, Knowledge, and Attitudes Relating to the Acquired Immunodeficiency Syndrome. *J. Gen. Intern. Med.* **3**(1988):161–165.
8. Kelly, J. A., Lawrence, J. S., Smith, S., Hood, H. V., and Cook, D. J. Medical Students' Attitudes towards AIDS and Homosexual Patients. *J. Med. Educ.* **62**(1987):549–556.
9. What Internists Think about AIDS. *IM: Intern. Med. for the Specialist.* **8**(October 15, 1987):53–60. [No author listed.]
10. Elford, J. Moral and Social Aspects of AIDS: A Medical Student's Project. *Soc. Sci. Med.* **24**(1987):543–549.

Willingness of Health-Professions Students to Treat Patients with AIDS

CHARLES J. CURREY, M.H.A., P.A.-C., MICHAEL JOHNSON, M.D., and
BARBARA OGDEN, M.S.N.

Abstract—This 1988–89 survey of 319 students in the medical, dental, nursing, and allied health-care professions revealed that over one-third had some reservations about treating acquired immunodeficiency syndrome (AIDS) patients. Most were unwilling to perform mouth-to-mouth resuscitation on patients with AIDS, and most also believed that health-care workers had the right to refuse care to AIDS patients. Unwillingness to treat AIDS patients was strongly associated with homophobic attitudes and concerns that patients with AIDS posed a risk to health professionals. AIDS education for health professionals should emphasize methods for the prevention of HIV infection among health workers, and include teaching strategies designed to deal with the irrational feelings that AIDS often engenders. *Acad. Med.* 65(1990):472–474.

Recent reports suggest that between 25 and 55% of nurses and physicians may be reluctant to work with patients with acquired immunodeficiency syndrome (AIDS) because of concerns over contagion.[1–4] In spite of the demonstrated low risk of occupational human immunodeficiency virus (HIV) infection among health-care workers,[5] some health-care personnel have refused to treat AIDS patients,[6] while others have experienced apprehension, reduced empathy for AIDS patients, and prejudicial attitudes toward homosexuals and drug abusers.[7–10] This paper is an empirical description of the attitudes of a group of students in the medical, dental, nursing, and allied health-care professions toward AIDS patients and their care.

Method

During the 1988–89 academic year, all first-year medical and dental students, junior-year nursing students, and junior-year physician assistant and physical therapy students training at a Southeastern university medical center treating AIDS patients were asked to anonymously complete a questionnaire dealing with their attitudes toward the care of AIDS patients and related issues. Because the medical and dental students completed the survey at the same time and did not distinguish their individual professions on the questionnaire, the two groups had to be combined for statistical purposes.

The questionnaire assessed the students' willingness to treat AIDS patients, their understanding of the basic epidemiology of AIDS, perceptions of the risk of AIDS transmission in the medical care environment, and attitudes about ethical issues surrounding AIDS care. Of 393 questionnaires, 319 were returned, for an overall response rate of 81.2%.

Results

As shown in Table 1, the majority of the students indicated that they would be willing to treat patients with AIDS. However, over one-third expressed some reservations—16.9% said they would not work with AIDS patients if given a choice, and 23.8% were uncertain about working with such patients if they had a choice. The percentages of those who were willing to work with AIDS patients were similar for all of the health professions sampled. Those willing and those unwilling to treat AIDS patients were also similar in age, gender, ethnic background, and marital status.

Virtually all the students understood the pathogenesis of AIDS and correctly identified the major risk factors and modes of transmission for the infection. However, less than half of those surveyed felt that they had learned enough about AIDS during their professional education to work safely with AIDS patients. The medical-dental student group appeared to be least comfortable with their AIDS knowledge. Among all the respondents, there was no statistically significant relationship between knowledge of AIDS epidemiology and willingness to treat AIDS patients.

A surprisingly large number of the students (21.6%) believed that health professionals were at high risk for AIDS. Most were inclined to take the precautions of wearing either gloves or a mask or both when performing procedures and providing routine care to patients with AIDS. Only 24.1% were willing to perform mouth-to-mouth resuscitation on AIDS patients.

Overall, most of the students (71.2%) believed that by working with AIDS patients they were placing their own health in jeopardy. Over half felt that health-care workers should have the right to refuse care to AIDS patients, and 66.8% believed that all patients (especially homosexuals and intravenous drug abusers) should be tested for AIDS when admitted to hospitals. Not surprisingly, those unwilling to treat AIDS patients were more likely to have these options. Those unwilling to treat AIDS patients were also three times more likely to believe that AIDS patients should be treated only in special hospitals and clinics.

Mr. Currey is assistant professor, the College of Health Related Professions; Dr. Johnson is clinical fellow, the Infectious Disease Division of the College of Medicine; and Ms. Ogden is assistant professor, the College of Nursing; all at the J. Hillis Miller Health Center, University of Florida, Gainesville.

Correspondence and requests for reprints should be addressed to Mr. Currey, Box J-176 JHMHC, University of Florida, Gainesville, FL 32610.

Other articles about the effects of AIDS on educating health-care providers appear on pages 446, 458, 461, 464, 467, 470, and 482.

Table 1

Percentage of 319 Health-Professions Students Who Answered "Yes" to Various Questions about Their Attitudes toward Treating Patients with AIDS, 1988–89[*]

Question	All Students ($n = 319$)	Medical-Dental[†] ($n = 151$)	Nurse ($n = 83$)	Physician Assistant ($n = 58$)	Physical Therapist ($n = 27$)	Those Willing to Treat AIDS Patients[‡] ($n = 190$)	Those Unwilling to Treat AIDS Patients[‡] ($n = 54$)
	Percentage of Students Answering "Yes"						
		Students Listed by Their Health Profession ($n = 319$)					
If given a choice, would you be willing to treat patients with AIDS?	59.6	51.7	67.5	74.1	48.1	100.0	0
If given a choice, would you be willing to treat homosexuals?	71.5	67.6	77.1	77.6	63.0	95.8	16.7[§]
If given a choice, would you be willing to treat IV drug abusers?	72.7	67.6	79.5	81.0	63.0	93.7	29.6[§]
Would you be willing to perform mouth-to-mouth resuscitation on an AIDS patient in respiratory arrest?	24.1	27.8	20.5	20.7	22.2	31.6	7.4[§]
Has your professional education provided you with enough information to work safely with AIDS patients?	46.1	24.5[§]	73.5	55.2	63.0	50.5	46.3
Are you concerned that working with AIDS patients may endanger your health?	71.2	76.8	66.3	63.8	70.4	60.0	98.2[§]
Should health-care workers have the right to refuse working with AIDS patients?	53.9	51.7	62.7	37.9[‖]	74.1	44.2	81.5[§]
Should hospitals be able to terminate employment of health professionals who refuse to work with AIDS patients?	19.4	19.9	15.7	25.9	14.8	24.7	3.7[§]
Should AIDS patients be treated only in special hospitals and clinics?	28.8	35.1	15.7	24.1	44.4	20.0	63.0[§]
Homosexuals who have AIDS deserve their fate because of their lifestyles.	10.0	9.3	12.1	10.3	7.4	6.3	25.9[§]
I am sympathetic to the fear and misery AIDS has caused the gay community	69.9	67.6	73.5	70.7	70.4	76.8	57.4[§]
Homosexuals should have the right to practice their sexual preferences as they choose.	69.9	69.5	72.3	72.4	59.3	78.4	40.7[§]

[*] In 1988–89, 319 students in the medical, dental, nursing, and allied health-care professions at a university medical center in the Southeast anonymously completed a questionnaire that elicited their attitudes toward the care of AIDS patients and related issues.

[†] The medical and dental students did not note their individual professions on the questionnaire; thus, the responses of these two groups have been combined.

[‡] Only 244 of the students gave a definite answer (that they either were willing or were unwilling to treat AIDS patients); the others answered "unsure."

[§] Statistically significant differences in "yes" responses among the students in the different health professions, and between those willing and unwilling to treat AIDS patients, chi-square comparison indicates that $p < .001$.

[‖] Statistically significant differences in "yes" responses among the students in the different health professions, and between those willing and unwilling to treat AIDS patients, chi-square comparison indicates that $p < .01$.

The majority of the students expressed compassion for the gay community's struggle with AIDS, and believed that homosexuals had the right to practice their sexual preferences as they chose. Only 10% felt that homosexuals with AIDS deserved their fate because of their lifestyles. However, those unwilling to treat AIDS were more likely to have had homophobic opinions than those willing to treat the disease.

Conclusions and Discussion

Students in this study had many of the same concerns about AIDS care reported by other health professionals.[1-4,7-10] In particular, they expressed fear of contagion and perceived themselves to be at increased risk when working with AIDS patients. Caution should be exercised in attempting to generalize from this study, however, since it did not involve a random sample, and because it is likely that students in other regions of the country—with more or less exposure to AIDS patients—would have different perceptions of risk.

Written comments provided by the students at the end of the questionnaire indicated that some of the anxiety over occupational HIV infection may have stemmed from perceived inconsistencies in the way information about AIDS was being presented in the classroom, and skepticism about what was being reported in the scientific literature about AIDS. For example, one student cited two different instructors quoting widely differing rates of HIV infection among health-care workers exposed to needlesticks. Another student viewed as contradictory the facts that the human immunodeficiency virus had been detected in saliva, but saliva was not capable of transmitting the infection.

It is alarming that only about 25% of the students surveyed would have been willing to perform mouth-to-mouth resuscitation on AIDS patients. However, written comments seemed to suggest that many experienced some conflict over this issue. While recognizing their obligation to intervene in an emergency to save a life, some nonetheless felt that this obligation was outweighed by the necessity not to hazard their own lives. Others rationalized their decision by questioning the wisdom of resuscitating a patient with an incurable terminal disease.

Overcoming irrational fear of AIDS among health-professions students will be easier if instructors keep information on AIDS current and avoid anxiety-provoking inconsistencies and "mixed messages" when presenting facts about the disease—especially facts dealing with the safety of working with AIDS patients. While students should be made aware of the dangers inherent in exposure to any infectious disease, they should also understand that adherence to procedures such as the Universal Precautions can significantly reduce risks of contagion when treating patients with AIDS.[11]

To help allay fear and modify attitudes about AIDS patients greater emphasis needs to be placed on ethics and human values in health-care curricula to help students prepare for the affective aspects of caring for AIDS patients. Discussion groups or role-playing exercises guided by experienced clinicians may help students vent fears of death and deal with the discomfort often associated with the topics of sexually transmitted diseases and homosexuality. It might also be instructive to invite volunteer AIDS patients and members of their families into the classroom to help underscore the humanity of AIDS patients.

References

1. Blumenfield, M., Smith, P. J., Milazzo, J., Seropian, S., and Wormser, G. P. Survey of Attitudes of Nurses Working With AIDS H., Freeman, K., and Shelov, S. P. Concerns of Medical and Pediatric House Of- Patients. *Gen. Hosp. Psych.* **9**(1987): 58-63.
2. Link, R. N., Feingold, A. R., Charap, M. ficers about Acquiring AIDS from Their Patients. *Am. J. Public Health* **78**(1988): 455-459.
3. Richardson, J. L., Lochner, T., McGuigan, K., and Levine, A. M. Physician Attitudes and Experience Regarding the Care of Patients with Acquired Immunodeficiency Syndrome and Related Disorders. *Med. Care* **25**(1987):675-685.
4. Imperato, P. J., Feldman, J. G., Nayeri, K., and Dehovitz, J. A. Medical Students' Attitudes Towards Caring for Patients with AIDS in a High Incidence Area. *NY State J. Med.* **88**(1988):223-227.
5. Marcus, R. Surveillance of Health Care Workers Exposed to Blood from Patients Infected with the Human Immunodeficiency Virus. *N. Engl. J. Med.* **319**(1988):1118-1123.
6. Gruson, L. AIDS Fear Spawns Ethics Debate as Some Doctors Withhold Care. *New York Times*, July 11, 1987, p 1.
7. Treiber, F. A., Shaw, D., and Malcolm, R. Acquired Immune Deficiency Syndrome: Psychological Impact on Health Personnel. *J. Nerv. Ment. Dis.* **175**(1987):496-499.
8. Douglas, C. J., Kalman, C. M., and Kalman, T. P. Homophobia Among Physicians and Nurses: An Empirical Study. *Hosp. Comm. Psych.* **36**(1985):1309-1311.
9. Kelly, J. A., St. Lawrence, J. S., Hood, H. V., Smith, S., and Cook. D. J. Nurses' Attitudes Towards AIDS. *J. Cont. Ed. Nurs.* **19**(1988):78-83.
10. Royse, D., and Birge, B. Homophobia and Attitudes Towards AIDS Patients among Medical, Nursing, and Paramedical Students. *Psychol. Rep.* **61**(1987):867-870.
11. Centers for Disease Control. Update: Universal Precautions for Prevention of Transmission of Human Immunodeficiency Virus, Hepatitis B Virus, and Other Blood-borne Pathogens in Health Care Settings. *MMWR* **37**(1988):377-388.

AAMC DATA REPORT

Graduating Medical Students' Debt and Specialty Choices

The average debt level of indebted respondents to the AAMC Graduation Questionnaire is reported here for each of the specialty career choices of senior medical students in 1986 and 1989. The reported debt combines premedical educational debt and medical school educational debt. Table 1 shows for each specialty the percentages of respondents with debts and their average debts in 1986 and 1989, as well as the percentage increase in average debt for indebted students over these years.

Between 1986 and 1989, the average debt level for indebted students rose from $33,499 to $42,374. This increase of 26.5% is more than twice the 13.1% inflation in the consumer price index over that time.

In 1986, seniors with specialty career choices in general pediatrics, ophthalmology, radiology, general and subspecialty internal medicine, pathology, physical medicine, general surgery, and family practice had the lowest average debt levels. In 1989, graduating students with specialty choices in general pediatrics, pathology, urology, family practice, general surgery,

Table 1

Average Debt and Percentage of Graduating Medical Students with Debt, and the Percentage of Increase in Average Debt, AAMC Graduation Questionnaire, 1986 and 1989

	1986		1989		
Specialty	Ave. Debt of Indebted Respondents	% with Debt	Ave. Debt of Indebted Respondents	% with Debt	Percent Increase in Ave. Debt, 1986-1989
Anesthesiology	34,241	86.1	41,901	84.0	+22.4
Dermatology	34,053	77.9	39,390	74.3	+15.6
Emergency medicine	37,637	86.9	48,709	83.3	+29.4
Family practice	33,326	84.6	41,005	84.8	+23.0
General internal medicine	31,630	82.4	43,355	82.8	+37.1
Internal medicine subspecialties*	31,916	80.8	43,602	82.5	+36.6
Neurological surgery	36,670	77.9	43,940	81.3	+19.8
Neurology-child neurology	36,176	75.7	45,144	79.3	+24.8
Obstetrics-gynecology and subspecialties†	34,918	84.1	45,757	81.7	+31.0
Ophthalmology	30,232	78.2	39,140	75.5	+29.5
Orthopedic surgery	34,700	81.2	42,340	79.1	+22.0
Otolaryngology	34,278	84.5	42,164	80.7	+23.0
Pathology	32,874	76.1	38,041	78.2	+15.7
General pediatrics	30,065	80.7	37,709	82.7	+25.4
Pediatrics subspecialties‡	35,643	78.9	44,427	82.9	+24.6
Physical medicine-rehabilitation	33,245	87.8	47,792	81.6	+43.8
Psychiatry-child psychiatry	35,576	84.7	43,521	83.9	+22.3
Radiology and subspecialties	30,983	83.4	43,614	80.6	+40.8
General surgery	33,268	80.1	41,681	80.1	+25.3
Surgical subspecialties§	38,766	76.5	46,162	78.8	+19.1
Urology	37,412	74.0	38,288	74.7	+2.3
All respondents	33,499	82.1	42,374	81.3	+26.5

*Internal medicine subspecialties are cardiovascular disease, critical care, endocrinology and metabolism, gastroenterology, hematology, infectious disease, medical oncology, nephrology, pulmonary disease, and rheumatology.

†Obstetrics-gynecology subspecialties are gynecologic oncology, maternal and fetal medicine, and reproductive endocrinology.

‡Pediatrics subspecialties are cardiology, critical care, endocrinology, hematology-oncology, nephrology, and neonatal-perinatal medicine.

§Surgical subspecialties are critical care, pediatric surgery, colon and rectal surgery, plastic surgery, and thoracic surgery.

and anesthesiology had the lowest average debt levels. The highest average debt levels of graduates in both years were for those selecting emergency medicine and surgical subspecialties. Compared with 1986, the average debt of indebted graduates selecting general and subspecialty internal medicine in 1989 was slightly larger than the overall average debt. In 1989, graduates with specialty choices in emergency medicine, physical medicine, surgical subspecialties, obstetrics-gynecology and its subspecialties, and neurology, had the highest average debt levels.

The percentage of indebted students dropped slightly from 82.1% to 81.3%, but the percentage changes differed by specialty. The percentages with debt increased for pediatric subspecialties, neurology, and neurological surgery, while the percentages decreased for physical medicine, otolaryngology, dermatology, and emergency medicine.

Randolph Park, Ph.D.

Dr. Park is a senior research associate, Section on Student and Educational Programs, Association of American Medical Colleges, Washington, D.C.

A3273-8

AUGUST 1990

ACADEMIC MEDICINE

JOURNAL OF THE ASSOCIATION OF AMERICAN MEDICAL COLLEGES

Medical Education in China

Management of Academic Medicine

Development of Medical Education

Medical Library Education

National Examination Validity

Performance on the ABIM Examination

Quarterly Profile Examination

College Health Services as Training Sites

Students' Research and Later Activities

Physicians' Patient Load and Resource Use

Consensus on Nutrition Teaching

National Policy Perspectives

VOLUME 65 • NUMBER EIGHT

AUGUST 1990
VOLUME 65
NUMBER EIGHT

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Special Articles:

China's Quest for Management of Academic Medicine, *by William T. Butler, M.D., and Steven J. Ruma, Ph.D.* 497

Medical Library Education in China, *by David S. Crawford, Dip.Lib.Stud., and Dizhi Xiong, M.D.* 502

Predictive Validity of a National Examination for Medical Graduates in the People's Republic of China, *by Liu Bingxun, M.D., Mei Renlang, M.D., Niu Weidong, M.D., Zhang Zu, M.D., Qui Xiangxing, M.D., and Wang Lie, M.D.* 505

International Medical Education

Development of Medical Education in China, *by Deng Yizhong, M.D.* 512

Ideas for Medical Education

The Quarterly Profile Examination, *by Louise Arnold, Ph.D., and T. Lee Willoughby, M.S.* 515

From the President 517

National Policy Perspectives 518

AAMC Focus 520

Commentary

Medical Research and Research!America, *by Lowell Weicker, Jr.* 521

Letter to the Editor

Revitalizing the Undergraduate College–Medical School Linkage, *by William A. Thomson, Ph.D., Leslie M. Miller, Ph.D., Quentin W. Smith, M.S., and Zenaido Camacho, Ph.D.* 522

Research Reports
ARTICLES

Performances of U.S. Osteopathic and Canadian Medical School Graduates on the American Board of Internal Medicine Certifying Examinations, 1984–1988, *by Judy A. Shea, Ph.D., John J. Norcini, Ph.D., and John A. Benson, Jr., M.D.* 523

College Health Services as a Site for Ambulatory Care Training of Residents and Medical Students, *by John M. Chuck, M.D., Robert C. Davidson, M.D., M.P.H., and Suzanne Snively, M.D.* 527

The Association between Students' Research Involvement in Medical School and Their Postgraduate Medical Activities, *by Scott Segal, M.D., Tom Lloyd, Ph.D., Peter S. Houts, Ph.D., Paula L. Stillman, M.D., Robert L. Jungas, Ph.D., and Robert B. Greer III, M.D.* 530

Physicians' Patient Load per DRG, the Consumption of Hospital Resources, and the Incentives of the DRG Prospective Payment System, *by Eric Munoz, M.D., M.B.A., Katherine Mulloy, Jonathan Goldstein, M.P.A., Jordon Josephson, M.D., Noel Tenenbaum, M.D. and Leslie Wise, M.D.* 533

SHORT PAPERS

Priorities for Nutrition Content within a Medical School Curriculum: A National Consensus of Medical Educators, *by the American Society for Clinical Nutrition Committee on Medical/Dental School and Residency Nutrition Education* 538

A Survey of the Personal Computer Laboratory Facilities and Policies at U.S. Medical Schools, *by Linda K. Gunzburger, Ph.D., and Thomas Loesch, M.Ed.* 540

Brief Communications

Amount and Distribution of Interns' Working Hours, *by Frank J. Kroboth, M.D., Alan R. Whitman, Michael Karpf, M.D., and Gerald S. Levey, M.D.* 542

Two Methods of Including Postgraduate Students as Observers of Psychotherapy, *by Hernan Davanzo, M.D., Claudia Pruzzo, M. Teresa Sanfuentes, Gabriela Trivelli, and Verónica Wigdorsky* 542

The Effect of Surgery Clerkship Timing on Fourth-Year Students' Surgery Knowledge, *by Frank A. Baciewicz, Jr., M.D., Judy Fagley, Michael Weaver, M.S.N., Richard Yeasting, Ph.D., and Neil R. Thomford, M.D.* 543

Erratum 543

Publishing Services: Hanley & Belfus, Inc.
210 South 13th Street, Philadelphia, PA 19107

AUGUST 1990
VOLUME 65
NUMBER EIGHT

ACADEMIC MEDICINE

JOURNAL OF THE ASSOCIATION OF AMERICAN MEDICAL COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
Edward J. Stemmler, M.D.
Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
University of Kentucky College of Medicine

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
Georgetown University Hospital
Philip Anderson, M.D.
University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
Harvard Medical School
Eta Berner, Ed.D.
University of Alabama School of Medicine
Ruth Bulger, Ph.D.
National Academy of Sciences
Vincent A. Fulginiti, M.D.
Tulane University School of Medicine
Nancy Gary, M.D.
Albany Medical College
Beatrix Hamburg, M.D.
Mt. Sinai Hospital
Murray M. Kappelman, M.D.
The University of Maryland School of Medicine
Ralph W. Muller
University of Chicago Hospitals
Herbert Pardes, M.D.
College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
University of Alabama School of Medicine
W. Vickery Stoughton
Toronto General Hospital
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
David H. Cohen, Ph.D.
Northwestern University

Chairman-Elect
William T. Butler, M.D.
Baylor College of Medicine

Immediate Past Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center School of Medicine

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Joe Dan Coulter, Ph.D.
University of Iowa College of Medicine

Chairman, Council of Teaching Hospitals
Raymond G. Schultze, M.D.
University of California, Los Angeles Medical Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

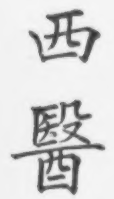
All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.
Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036; (202) 828-0416.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

LIU BINGXUN, M.D., MEI RENLANG, M.D., NIU WEIDONG, M.D.,
ZHANG ZU, M.D., QIU XIANGXING, M.D., and WANG LIE, M.D.

Predictive Validity of a National Examination for Medical Graduates in the People's Republic of China

Abstract—National examinations for medical graduates were introduced on an experimental basis in the People's Republic of China in 1982. To estimate the predictive validity of the National Medical Examination (NME), an investigation of the postgraduate competence of a sample of the participating examinees was conducted in 1984. The sample consisted of 1,717 of the 4,995 graduates from 13 medical colleges who had taken the initial NME. Their scores on the NME and the ratings given them by directors of postgraduate programs in nine aspects of clinical competence were compared by frequency distribution and product-moment correlation coefficients. Scores on the NME were consistent with measures of postgraduate clinical competence and, as a whole, correlated significantly with the ratings of clinical competence, supporting the use of the score on the NME as a predictor of postgraduate clinical competence. However, the extent of the relationship between the NME score and postgraduate clinical competence varied according to the specialty program of postgraduate medical training. *Acad. Med.* 65(1990):505–511.



To accumulate the experience needed to establish a system of national medical qualifying examinations and to evaluate objectively the educational quality of the medical schools in the People's Republic of China, the Ministry of Public Health has administered National Medical Examinations (NMEs) to graduates of some of these medical schools annually since 1982.

The NMEs have each been composed of 200 to 300 multiple-choice questions about internal medicine, surgery, obstetrics-gynecology, and pediatrics, to be completed in six hours. Test items are selected from the Test Item Bank developed by the Committee of Test Item Formulation for the National Medical Examination.

When the first NME was given, in 1982, 4,995 medical graduates from 13 medical colleges were asked to take this examination. Two hundred test items were put to the examinees; their average score was 169.5. The initial results of this examination were analyzed and found to be content-valid and reliable, with a Kuder-Richardson (KR-20) reliability coefficient of 0.78.[1] A second study, of 226 examinees, demonstrated that the product-moment correlations between scores on the NME and undergraduate performances in basic and clinical sciences in the medical school were 0.48 and 0.43, respectively.[2] Two other studies demonstrated reliability coefficients of 0.58 and 0.65, and correlations with undergraduate performances in medical schools ranging from 0.26 to 0.69.[3,4]

Whether the results of such an examination can be used as a predictor of the postgraduate clinical competence of the examinee has been debated. Kappelman[5] argued that multiple-choice questions could be used to measure examinees' mastery of knowledge, but could not be used to assess such clinical abilities as appropriate interaction with patients, skill in accessing the literature, and technical skills. Various investigators[6–8] have studied the predictive validity of qualifying examinations. Some concluded that an important relationship existed between the scores on a nationally standardized examination and ratings of the postgraduate clinical competence of the examinees, while others reached an utterly contrary conclusion. Gonnella and Veloski and their colleagues[6,7] reported a positive association between scores on the examinations of the National Board of Medical Examiners (NBME) and ratings of clinical competence at the end of the first postgraduate year; however, the predictive validity of the NBME examination scores depended upon the types of specialty training programs that the examinees were taking.[6,7] In order to ascertain the predictive validity of the first NME, we investigated the postgraduate clinical competence of a sample of the examinees in the fall of 1984.

Method

The sample for the present study consisted of 1,717 residents who graduated from 13 medical schools and took the first NME in 1982. Of the total group, 799 were training in internal medicine, 499 in surgery, 260 in obstetrics-gynecology, and 159 in pediatrics. They constituted 34.4% of the 4,995 examinees taking the NME in that year. The other 3,278 medical

---

Dr. Liu Bingxun is director, Department of Medical Education, and chairman of the Committee on Test Item Formulation for National Medical Examination, Ministry of Public Health, Beijing, People's Republic of China; Dr. Mei Renlang is associate professor of medical education, Shanghai Medical University; Dr. Niu Weidong is director, Division of Research and Statistics, National Medical Examination Center; Dr. Zhang Zu is research assistant, Division of Medical Education Research, Norman Bethune Medical University; Dr. Qiu Xiangxing is director, Division of Medical Education, Shanghai Second Medical University; and Dr. Wang Lie is research assistant, Medical Education Research Center, China Medical University.

Correspondence should be addressed to Dr. Liu, Director, Department of Medical Education, Ministry of Public Health, People's Republic of China, 44 Houhai Beiyan, Beijing, China. Reprints will not be available.

The message of the calligraphy is "Chinese traditional medicine."

graduates were taking postgraduate training in basic medical sciences or other clinical specialty programs, or were engaged in other activities, and were therefore excluded from the study.

Rating forms (one for each resident) were sent to the residency supervisors, who were asked to rate the clinical competences of their trainees on a four-point scale ranging from 1 for unsatisfactory to 4 for excellent. The form included 44 items that described nine aspects of clinical competence: data gathering, physical examination, differential diagnosis, patient management, clinical procedures, application of basic science knowledge, verbal communication, service attitude, and comprehensive assessment. ("Comprehensive assessment" was not the sum or the average value of the ratings of the other eight aspects, but a general evaluation of a resident based on his or her clinical performance as a whole and the general impression on the rater.) The average value of the ratings of all items on each aspect of competence was computed for each resident; the residents were then classified accordingly into three ability groups: those whose average ratings on a particular aspect were in the top 15% were put into the high-ability group for that aspect, while those in the lowest 15% were considered to be in the low-ability group. The remaining residents constituted the medium-ability group.

In the same way, according to their scores on the NME, the residents were again divided into three groups: high-score group, medium-score group, and low-score group. The two sets of data were compared statistically by their joint frequency distributions and Pearson product-moment correlation coefficients.

Results

Table 1 shows (1) the numbers of graduates classified by their NME scores (low, medium, or high) and their postgraduate clinical competence ratings (low-, medium-, or high-ability) for each of the nine aspects of clinical competence (the count); (2) the percentage distributions of the postgraduate competence ratings, by the NME scores (the column percentage); (3) the percentage distributions of the NME scores, by postgraduate competence ratings (the row percentage); and (4) chi-square (χ^2) for each small data matrix.

Column Percentages

An analysis of the percentage distributions of the levels of ratings of postgraduate clinical competence

Table 1: Numbers of Graduates and Percentage Distributions of National Medical

Aspect of Competence	Rating	No. of Graduates (n = 1,717)		Total			
				High (n = 271)	Med. (n = 1,163)	Low (n = 283)	χ^2
Comprehensive assessment	High	261	Count	50	192	19	42.322*
			Col Pct	18	17	7	
			Row Pct	**19**	**74**	**7**	
	Med.	1,211	Count	199	815	197	
			Col Pct	73	70	70	
			Row Pct	**16**	**67**	**16**	
	Low	245	Count	22	156	67	
			Col Pct	8	13	24	
			Row Pct	**9**	**64**	**27**	
Data gathering	High	828	Count	144	585	99	29.902*
			Col Pct	53	50	35	
			Row Pct	**17**	**71**	**12**	
	Med.	579	Count	88	383	108	
			Col Pct	32	33	38	
			Row Pct	**15**	**66**	**19**	
	Low	310	Count	39	195	76	
			Col Pct	14	17	27	
			Row Pct	**13**	**63**	**25**	
Physical examination	High	639	Count	113	452	74	37.495*
			Col Pct	42	39	26	
			Row Pct	**18**	**71**	**12**	
	Med.	489	Count	82	338	69	
			Col Pct	30	29	24	
			Row Pct	**17**	**69**	**14**	
	Low	589	Count	76	373	140	
			Col Pct	28	32	49	
			Row Pct	**13**	**63**	**24**	

within different levels of NME scores demonstrated that the two were consistent with each other in most aspects; the exceptions were data gathering and differential diagnosis (see the column percentage distribution in Table 1).

The proportion of graduates with high NME scores was comparatively greater in the groups with high clinical ability than in the groups with low clinical ability. In addition, the proportion of graduates with low NME scores was comparatively larger in the groups with low clinical competency ratings than in those with high ratings. So with regard to comprehensive assessment of clinical competence, 18% of the residents in the group with high total scores were identified as members of the high-ability group, but only 8% were rated as having low ability. Conversely, only 7% of the residents in the group with low total scores were classified into the group with high ability, while up to 24% were rated as having low ability.

However, with respect to data-gathering skills, the proportion of graduates with low NME scores was comparatively larger in the group with high ability than in the group with low ability. Thirty-five percent of the residents in the low-total-score group were considered to have high ability, while only 27% of them were put into the low-ability group. Conversely, only 26% of the residents with high total scores were in the high-ability group for differential diagnosis, while the corresponding percentage of those with high total scores and in the low-ability group was as high as 40%. The same phenomenon was observed in the relationships between the ratings on these two competence aspects and the NME scores in various disciplines.

Row Percentages

Table 1 also displays the relationships among the variables of postgraduate competence and NME scores in another way — by row percentage distribution, that is, the percentage distribution of NME scores by postgraduate competence ratings. The ratings of clinical competence after graduation tended to be consistent with the NME scores in all the nine aspects, including data gathering and differential diagnosis. For example, in the comprehensive assessment area, 19% of the graduates in the high-ability group achieved high total scores on the NME, but only 7% of them belonged to the group with low total NME scores. At the same time, only 9% of the graduates with low ability ratings achieved good performances on the NME, while the percentage in the group who had low NME

Examination Scores and Postgraduate Competence Ratings, People's Republic of China, 1984

National Medical Examination Score															
Int. Med. Component				Surgery Component				Ob.-Gyn. Component				Pediatrics Component			
High (n = 347)	Med. (n = 1,058)	Low (n = 312)	χ^2	High (n = 307)	Med. (n = 1,098)	Low (n = 312)	χ^2	High (n = 267)	Med. (n = 1,140)	Low (n = 310)	χ^2	High (n = 488)	Med. (n = 880)	Low (n = 349)	χ^2
59	169	33	25.210*	61	172	28	26.733*	48	172	41	5.080	82	146	33	17.538*
17	16	11		20	16	9		18	15	13		17	17	9	
23	**65**	**13**		**23**	**66**	**11**		**18**	**66**	**16**		**31**	**56**	**13**	
254	747	210		214	779	218		184	812	215		347	615	249	
73	71	67		70	71	70		69	71	69		71	70	71	
21	**62**	**17**		**18**	**64**	**18**		**15**	**67**	**18**		**29**	**51**	**21**	
34	142	69		32	147	66		35	156	54		59	119	67	
10	13	22		10	13	21		13	14	17		12	14	19	
14	**58**	**28**		**13**	**60**	**27**		**14**	**64**	**22**		**24**	**49**	**27**	
187	509	132	14.075*	169	535	124	17.105*	150	549	129	13.041†	252	431	145	11.729*
54	48	42		55	49	40		56	48	42		52	49	42	
23	**61**	**16**		**20**	**65**	**15**		**18**	**66**	**16**		**30**	**52**	**18**	
110	364	105		88	376	115		74	391	114		156	300	123	
32	34	34		29	34	37		28	34	37		32	34	35	
19	**63**	**18**		**15**	**65**	**20**		**13**	**68**	**20**		**27**	**52**	**21**	
50	185	75		50	187	73		43	200	67		80	149	81	
14	17	24		16	17	23		16	18	22		16	17	23	
16	**60**	**24**		**16**	**60**	**24**		**14**	**65**	**22**		**26**	**48**	**26**	
135	408	96	15.890*	126	424	89	26.516*	113	424	102	15.933*	210	319	110	19.335*
39	39	31		41	39	29		42	37	33		43	36	32	
21	**64**	**15**		**20**	**66**	**14**		**18**	**66**	**16**		**33**	**50**	**17**	
109	299	81		89	322	78		88	317	84		134	264	91	
31	28	26		29	29	25		33	28	27		27	30	26	
22	**61**	**17**		**18**	**66**	**16**		**18**	**65**	**17**		**27**	**54**	**19**	
103	351	135		92	352	145		66	399	124		144	297	148	
30	33	43		30	32	46		25	35	40		39	34	42	
17	**60**	**23**		**16**	**60**	**25**		**11**	**68**	**21**		**24**	**50**	**25**	

Continued —

Table 1: Numbers of Graduates and Percentage Distributions of National Medical

Aspect of Competence	Rating	No. of Graduates (n = 1,717)		Total			χ^2
				High (n = 271)	Med. (n = 1,163)	Low (n = 283)	
Differential diagnosis	High	434	Count	72	317	45	20.341*
			Col Pct	26	27	16	
			Row Pct	**17**	**73**	**10**	
	Med.	528	Count	91	361	86	
			Col Pct	34	30	30	
			Row Pct	**17**	**66**	**16**	
	Low	755	Count	108	495	152	
			Col Pct	40	43	54	
			Row Pct	**14**	**66**	**20**	
Patient management	High	318	Count	60	231	27	39.251*
			Col Pct	22	20	10	
			Row Pct	**19**	**73**	**8**	
	Med.	1,159	Count	184	786	189	
			Col Pct	68	68	67	
			Row Pct	**16**	**68**	**16**	
	Low	240	Count	27	146	67	
			Col Pct	10	13	24	
			Row Pct	**11**	**61**	**28**	
Clinical procedures	High	315	Count	58	225	32	25.286*
			Col Pct	21	19	11	
			Row Pct	**20**	**71**	**9**	
	Med.	1,119	Count	176	763	180	
			Col Pct	65	66	63	
			Row Pct	**17**	**68**	**15**	
	Low	283	Count	37	175	71	
			Col Pct	14	15	25	
			Row Pct	**15**	**62**	**23**	
Application of basic science knowledge	High	316	Count	68	222	26	43.255*
			Col Pct	25	19	9	
			Row Pct	**22**	**70**	**8**	
	Med.	1,130	Count	178	766	186	
			Col Pct	66	66	66	
			Row Pct	**16**	**68**	**16**	
	Low	271	Count	25	175	71	
			Col Pct	9	15	25	
			Row Pct	**9**	**65**	**26**	
Verbal communication	High	460	Count	76	327	57	28.986*
			Col Pct	28	28	20	
			Row Pct	**17**	**71**	**12**	
	Med.	1,004	Count	169	677	158	
			Col Pct	62	58	56	
			Row Pct	**17**	**67**	**16**	
	Low	253	Count	26	159	68	
			Col Pct	10	14	24	
			Row Pct	**10**	**63**	**27**	
Service attitude	High	597	Count	112	408	77	25.785*
			Col Pct	41	35	27	
			Row Pct	**19**	**68**	**13**	
	Med.	708	Count	110	489	109	
			Col Pct	41	42	39	
			Row Pct	**16**	**69**	**15**	
	Low	412	Count	49	266	97	
			Col Pct	18	23	34	
			Row Pct	**12**	**65**	**24**	

*$p = .01$ for values of 13.277 or higher.
†$p = .05$ for values of 9.488 or higher.

Examination Scores and Postgraduate Competence Ratings, People's Republic of China, 1984—*Continued*

National Medical Examination Score

Int. Med. Component				Surgery Component				Ob.-Gyn. Component				Pediatrics Component			
High (n = 347)	Med. (n = 1,058)	Low (n = 312)	χ^2	High (n = 307)	Med. (n = 1,098)	Low (n = 312)	χ^2	High (n = 267)	Med. (n = 1,140)	Low (n = 310)	χ^2	High (n = 488)	Med. (n = 880)	Low (n = 349)	χ^2
88	286	60	12.360†	84	290	60	13.540*	73	285	76	7.210	134	227	73	7.478
25	27	19		27	26	19		27	25	25		27	26	21	
20	**66**	**14**		**19**	**67**	**14**		**17**	**66**	**18**		**31**	**52**	**17**	
121	311	96		95	346	87		95	346	87		158	263	107	
35	29	31		31	32	28		35	30	28		32	30	31	
23	**59**	**18**		**18**	**66**	**16**		**18**	**66**	**16**		**30**	**50**	**20**	
138	461	156		128	462	165		99	509	147		196	390	169	
40	44	50		42	42	53		37	45	47		40	44	48	
18	**61**	**21**		**17**	**61**	**22**		**13**	**68**	**19**		**26**	**52**	**22**	
72	205	41	27.190*	74	206	38	22.305*	60	212	46	9.075	103	164	51	17.312*
21	19	13		24	19	12		22	19	15		21	19	15	
23	**64**	**13**		**23**	**65**	**12**		**19**	**67**	**14**		**32**	**52**	**16**	
244	712	203		196	751	212		178	771	210		335	594	230	
70	67	65		64	68	68		67	68	68		69	68	66	
21	**61**	**18**		**17**	**65**	**18**		**15**	**67**	**18**		**29**	**51**	**20**	
31	141	68		37	141	62		29	157	54		50	122	68	
9	13	22		12	13	20		11	14	17		10	14	19	
13	**59**	**28**		**15**	**59**	**26**		**12**	**65**	**23**		**21**	**51**	**28**	
67	205	43	8.951	72	208	35	25.064*	62	201	52	15.510*	100	160	55	9.605†
19	19	14		22	19	11		23	17	17		20	18	16	
22	**65**	**13**		**23**	**66**	**10**		**22**	**62**	**16**		**32**	**51**	**17**	
229	686	204		199	714	206		174	756	189		327	562	230	
66	65	65		65	65	66		65	66	61		67	64	66	
21	**62**	**17**		**19**	**64**	**17**		**17**	**67**	**16**		**30**	**50**	**20**	
51	167	65		36	176	71		31	183	69		61	158	64	
15	16	21		12	16	23		12	16	22		13	18	18	
18	**60**	**21**		**15**	**62**	**23**		**13**	**65**	**23**		**23**	**57**	**21**	
70	208	38	24.279*	71	210	35	31.377*	63	208	45	8.865	112	160	44	25.395*
20	20	12		23	19	11		24	18	15		23	18	13	
22	**66**	**12**		**22**	**66**	**11**		**20**	**66**	**14**		**35**	**51**	**14**	
240	687	203		202	726	202		168	753	209		322	577	231	
69	65	65		66	66	65		63	66	67		66	66	66	
21	**61**	**18**		**18**	**64**	**18**		**15**	**67**	**18**		**28**	**51**	**20**	
37	163	71		34	162	75		36	179	56		54	143	74	
11	15	23		11	15	24		13	16	18		11	16	21	
14	**60**	**26**		**13**	**60**	**28**		**13**	**66**	**21**		**20**	**53**	**27**	
95	297	68	23.683*	89	299	72	14.275*	86	295	79	9.154	131	254	75	15.876*
27	28	22		29	27	23		32	26	25		27	29	21	
21	**65**	**15**		**19**	**65**	**16**		**19**	**64**	**17**		**28**	**55**	**16**	
216	615	173		188	639	177		154	673	177		288	514	202	
62	58	56		61	58	57		58	59	57		59	58	58	
22	**61**	**17**		**19**	**64**	**18**		**15**	**67**	**18**		**29**	**51**	**20**	
36	146	71		30	160	63		27	172	54		69	112	72	
10	14	23		10	15	20		10	15	17		14	13	21	
14	**58**	**28**		**12**	**63**	**25**		**11**	**68**	**21**		**27**	**44**	**28**	
138	364	95	14.876*	114	395	88	15.512*	100	394	103	8.254	189	309	99	17.447*
40	34	30		37	36	28		37	35	33		39	35	28	
23	**61**	**16**		**19**	**66**	**15**		**17**	**66**	**17**		**32**	**52**	**17**	
142	446	120		128	456	124		115	477	116		196	372	140	
41	42	38		42	42	40		43	42	38		40	42	40	
20	**63**	**17**		**18**	**64**	**18**		**16**	**67**	**17**		**28**	**53**	**20**	
67	248	97		65	247	100		52	269	91		103	199	110	
19	23	31		21	22	32		19	24	29		21	23	32	
16	*60*	*24*		*16*	*60*	*24*		*13*	*65*	*22*		*25*	*48*	*27*	

total scores was as high as 27%. The relationships between the scores in the various disciplines covered in the NME and the postgraduate clinical competence ratings were in the same direction.

NME Predictive Validity

In order to identify the predictive validity of the NME, the Pearson correlation coefficients between the scores on the NME and the ratings in each competence aspect were calculated. It was found that the NME total scores correlated with the ratings of clinical competence in all the nine aspects. The coefficients ranged from 0.1205 to 0.1946 and were statistically significant ($p < .01$). Generally, statistically significant correlations were also found between the NME total scores in the various specialty components and the ratings of the various clinical competences, although the correlations among a few variables, in obstetrics-gynecology in particular, were of lesser magnitude.

In exploring the NME's predictive validity, the individual specialty programs that the residents were pursuing were taken into consideration. Significant correlations were found among the variables for the group of residents training in internal medicine programs, except for the correlation between the NME score on obstetrics-gynecology and competence in physical examination. However, the correlations in the other groups did not follow this pattern. For surgery and obstetrics-gynecology, half of these correlation coefficients were not statistically significant, including those relating to the score of the same discipline in the NME. No significant correlation was found among the variables in the pediatrics group except that between the NME score in pediatrics and the rating of competence in making a diagnosis.

Discussion

It is important to clarify the relationship between NME scores and postgraduate clinical competence in assessing the validity of this measuring instrument. In educational measurement, it is desirable that a valid and reliable examination not only determine the current level of performance of the examinees but also have predictive value regarding the development of their professional careers. Since the concurrent validity and reliability of the NME had been confirmed in 1982, the authors hypothesized that those graduates who scored high on the examination should receive high clinical competence ratings and vice versa.

To test the hypothesis, two statistical approaches were used. The data presented in Table 1 provide evidence supporting the hypothesis. In addition, significant product-moment correlations were found among the various aspects of the NME total score and the ratings of postgraduate clinical competence. These findings are consistent with other reports.[7] Therefore, generally speaking, it seems that the NME scores can be useful in predicting postgraduate competence.

As shown in the column percentage distribution in Table 1, the ratings of two aspects of competence—data gathering and differential diagnosis—appear to have been inconsistent with NME scores. "Low scores but high ability" and "high scores but poor ability" were found in these two areas. But these phenomena may have been caused by differences in the criteria used in rating. Raters had lower standards for the aspect of data gathering, so more residents were rated into the high-ability group; this interpretation is supported by negative skewness in the distribution. Conversely, higher standards obtained for differential diagnosis; thus, more residents were rated as having low ability; this is supported by positive skewness in the distribution.

Comparison of the percentage distributions of the examinees in the high-score group and those in the low-score group in these two aspects shows that more examinees in the low-score group were really rated as having high ability than were rated as having low ability in data gathering (35% versus 27%). The former is 1.3 times as large as the latter. But the percentage of examinees rated as having high ability in the high-score group is even greater (53%) and is 3.8 times as large as the percentage for those rated as having low ability (14%) in this aspect.

Similarly, in the ratings of differential diagnosis, more examinees were classified in the group of low ability than in the group of high ability in the high-score group (40% versus 26%), and the former is 1.5 times as large as the latter. But the percentage of examinees rated as having low ability in the low-score group is even larger (54%) and 3.4 times as large as the percentage for those rated as having high ability (16%) in this area.

These false appearances can be clarified by another presentation shown in the row percentage distribution in Table 1, i.e., the percentage distribution of NME scores by postgraduate competence ratings. The table shows that in the data-gathering and differential diagnosis aspects, as in other aspects, the NME scores and competence ratings were in the same direction. There were more examinees with high NME scores in the high-ability groups, and the percentages are 17% versus 12% and 17% versus 10%, respectively. On the contrary, in the groups that had low ability, the proportions of examinees with high scores were smaller than those of examinees with low scores; the percentages are 13% versus 25% and 14% versus 20%.

In analyzing the relationships of the NME specialty component scores to the measures of postgraduate competence in different specialty training programs, we found a significant correlation for internal medicine, while the ratings on half of the clinical competence aspects in surgery and in obstetrics-gynecology did not significantly correlate with the NME scores. No significant correlation was found among almost all the variables in the pediatrics programs. The reasons for these findings may be as follows:

1. They may have occurred only by chance. It seems that chance is unlikely, however, because Gonnella's

study[6] of the relationship between school performance and postgraduate competence resulted in a similar pattern.

2. The sample may not have been large enough. Statistically, the larger the sample, the more likely that the correlation will be significant. In this study, the highest correlation was found among the variables in the internal medicine group, which happened to have the largest sample size compared with the other groups. For surgery, obstetrics-gynecology, and pediatrics, the levels of significance of the correlations among the variables increased with increasing sample size.

3. Different departments and raters may have had different standards even though they had been given a uniform checklist. The Shanghai First Medical College analyzed the consistency of ratings among different medical units and found that there were significant differences among them, which probably explains the variation in results. To improve inter-rater reliability, it seems that more detailed and concrete rating standards should be developed and the number of raters for each resident should be increased.

4. Less variation and a narrower range in grouping the NME scores and the measures of clinical competence create more homogeneity in the residents' groups. When distinctions are not clear enough, especially in a small sample, correlation coefficients will be quite small. If this is the case, other evaluation indexes should be considered in assessing the predictive validity of the NME.

Conclusions

The results of the present study confirm that there is a certain consistency between the NME scores and postgraduate clinical competence, suggesting that the score on the NME can be used as a predictor of postgraduate clinical competence. However, the study has also shown that the extent of the relationship between the NME score and postgraduate clinical competence varies with different specialty programs of postgraduate medical training.

References

*1. The Office of National Medical Examination, Ministry of Public Health, People's Republic of China. *The Report on a Tentative National Examination for Medical Graduates from 13 Medical Schools in 1982.* 1983.

*2. Fang Chunwang, et al. Analyses of the Results of the National Examination for Medical Graduates from Shandong Medical College in 1982. *Med. Educ.* **2**(1983):11-21.

*3. Mei Renlang, et al. *Analysis of the Reliability and Validity of the National Examination.* The Symposium on Medical Education, Shanghai First Medical College, 1983.

*4. Qiu Xiangxing, et al. *The Analysis of the Scores of the Medical Graduates in 1982 National Examination.* The Symposium on Medical Education, Shanghai Second Medical College, 1983.

5. Kappelman, M. M. The Impact of External Examinations on Medical Education Programs and Students. *J. Med. Educ.* **58**(1983):300-308.

6. Gonnella, J. S., and Hojat, M. Relationship Between Performance in Medical School and Postgraduate Competence. *J. Med. Educ.* **58**(1983):679-685.

7. Veloski, J., Herman, M. W., Gonnella, J. S., Zeleznik, C., and Kellow, W. F. Relationships Between Performance in Medical School and First Postgraduate Year. *J. Med. Educ.* **54**(1979):909-916.

8. Hubbard, J. P. *Measuring Medical Education.* 2d ed. Philadelphia, Pennsylvania: Lea and Febiger, 1979.

---

*Printed in Chinese

RESEARCH REPORTS

ARTICLES

Performances of U.S. Osteopathic and Canadian Medical School Graduates on the American Board of Internal Medicine Certifying Examinations, 1984–1988

JUDY A. SHEA, Ph.D., JOHN J. NORCINI, Ph.D., and JOHN A. BENSON, Jr., M.D.

Abstract—The performances of Canadian medical school graduates and U.S. osteopathic medical school graduates who first took the American Board of Internal Medicine Certifying Examination between 1984 and 1988 were compared with the performances, during the same period, of U.S. and foreign medical school graduates. Approximately 100 Canadian graduates took the examination each year; by 1988 the number of osteopathic graduates taking it was 102, double the number participating in 1984. Nearly all the Canadian graduates attended university or university-affiliated residencies, whereas half of the osteopathic graduates attended non-university-affiliated programs. For their overall clinical competence and for the eight components of clinical competence, the Canadian graduates were rated highest, followed by the U.S., osteopathic, and foreign graduates. The Canadians' average examination scores were also highest. The authors discuss the relatively low level of performance of the osteopathic graduates, but conclude that these graduates appear to be an untapped source of talented physicians for internal medicine residencies. The limitations of studying self-selected groups of candidates are also discussed. *Acad. Med.* 65(1990): 523–526.

Several studies have shown that physicians' performances during residency and at the completion of training are related to the types of medical school experiences they have had.[1-3] For example, graduates of medical schools in the United States (USMGs) consistently score higher than do graduates of foreign medical schools (FMGs) on the American Board of Internal Medicine (ABIM) certifying examination.[2-5] The performance differentials between USMGs and FMGs are also observed for evaluations of clinical competence given by program directors at the conclusion of residency training. Breakdowns of data for subgroups of FMGs show finer distinctions in performance. For example, citizens of the United States who attended a medical school in the Caribbean had lower scores and ratings of competence, in general, than did both foreign citizens attending foreign schools and U.S. citizens attending foreign schools in locations other than the Caribbean.[3]

Other studies have shown that performance on the ABIM certifying examination is related to the quality of the residency program, regardless of where one has attended medical school.[6-8] The advantage of attending a residency viewed as outstanding by experts in the field was observed for men and women who were first-time takers of the certifying examination and were from the combined group of United States and Canadian medical school graduates.[9] In addition, foreign medical graduates who had attended a university residency program did better on the certifying examination than did their peers who had attended a non-university-affiliated program.[2]

Overall, past studies provide data that guide expectations about groups of candidates seeking certification in internal medicine. Such information may be helpful to program directors and other professionals involved in the graduate medical education process. In order to more completely describe the relationship between type of undergraduate medical school and later performance, it would be useful to have information about two previously unstudied subgroups of candidates: graduates of Canadian medical schools and graduates of osteopathic medical schools. (All osteopathic medical schools are located in the United States.)

Each year, both Canadian and osteopathic medical schools graduate hundreds of physicians. In 1987, 1,763 physicians completed undergraduate medical training in Canada, and 1,579 received degrees in osteopathic medicine in the United States.[10,11] Available data suggest that sizeable numbers of graduates from both backgrounds take graduate

---

Dr. Shea is senior psychometrician, Dr. Norcini is executive vice-president for evaluation and research, and Dr. Benson is president at the American Board of Internal Medicine. Drs. Shea and Norcini are in Philadelphia, Pennsylvania, and Dr. Benson is in Portland, Oregon.

Correspondence and requests for reprints should be addressed to Dr. Shea, American Board of Internal Medicine, 3624 Market Street, Philadelphia, PA 19104.

training in internal medicine. As of September 1, 1987, a total of 436 osteopathic graduates were in internal medicine residency programs approved by the Accreditation Council for Graduate Medical Education.[12] In the same year, 1,256 residents were in training in internal medicine in a Canadian program, although the proportion of Canadian graduates is unreported.[10] Further study is needed because these data neither include Canadian graduates training in residency programs in the United States nor allow for the fact that not all trainees attempt voluntary certification.

This study details the performances of two groups of candidates taking the ABIM certifying examination for the first time between 1984 and 1988: graduates of Canadian medical schools and graduates of U.S. osteopathic medical schools. Numbers and types of residency programs attended are reported, and differences between program directors' rating of competence and examination performances are explored. Trends in numbers and performances are examined relative to those of two large baseline groups, USMGs and FMGs. Finally, the paper reports the joint effects of undergraduate training, type of residency program, and program director's rating of overall clinical competence on examination performance.

Method

This section describes the four groups studied and each measure on which they were compared.

Candidates

The candidates included in this study took the ABIM certifying examination in general internal medicine during the period 1984–1988. The study cohort in each year was limited to all first-time takers who had both finished their residencies and been evaluated by their program directors in the year of the examination. All the candidates had satisfactorily completed the equivalent of three years of training in a residency program accredited by the Accreditation Council for Graduate Medical Education (ACGME) or, if training had been completed in Canada, accredited either by The Royal College of Physicians and Surgeons of Canada or by the Professional Corporation of Physicians of Quebec. In addition, candidates with a Canadian undergraduate education must have attended a medical school accredited by the Committee for Accreditation of Canadian Medical Schools. Candidates from osteopathic medical schools must have graduated from colleges of osteopathic medicine accredited by the American Osteopathic Association.[13]

Types of Training

Residency programs were classified as (1) "top" U.S. programs (chosen by ballot of five members of the American Board of Internal Medicine, who were given a list of all accredited residencies), (2) university programs, and (3) non-university-affiliated programs. The second category consisted primarily of university and university-affiliated programs, although a small number of federal and military programs were included. The common feature of the final category was that all programs were independent of formal affiliations with universities. They were diverse, however, in that programs ranged from small hospitals training only a few residents per year to large programs that were well known and established.

Directors' Ratings

At the end of training, directors of internal medicine programs forwarded to the Board ratings of clinical competence for their residents who applied for the examination. Ratings were submitted for an overall evaluation of clinical competence and eight of its components: clinical judgment, medical knowledge, history-taking skill, physical examination skill, procedural skill, medical care, professional attitudes and behavior, and humanistic qualities. These evaluations were gathered on a nine-point scale where ratings of 1 to 3 were "unsatisfactory," 4 to 6 were "satisfactory," and 7 to 9 were "superior." Candidates who received overall ratings of 1 to 3 were not permitted to take the examination.

Examinations

The two-day written examinations were composed of patient management problems (in 1984–1986 only) and multiple-choice questions consisting of single-best-answer, matching (in 1984–1986 only), and multiple true-false items. A standardized score was developed for each reference group examinee for each item format. The reference group included all candidates who had graduated from U.S. or Canadian medical schools, were first-time takers of the certifying examination, and had recently completed three years of training in an approved residency program. Nonreference group members were given scores by locating their raw scores for each item format in the reference group distribution and assigning the associated standardized score. The standardized scores for each format were then weighted to form a composite which was restandardized to the reference group mean of 500 and a standard deviation of 100. This approach maintained the original rank ordering of candidates, as it placed all scores on the same scale. Additional details of the scoring process and estimates of reliability and validity can be found elsewhere.[6]

Results

In this section, tests of statistical significance are not provided. These data are based on the entire population of interest; consequently, there is no sampling error.

Numbers of Candidates

The number of Canadian graduates remained fairly stable, at about 100 candidates per year, except for a decrease to 72 in 1986. The number of osteopathic graduates nearly doubled over the five years, steadily increasing from 53 in 1984 to 102 in 1988. In comparison, the number of USMGs ranged from 3,435 to 3,659 and the number of FMGs ranged from 859 to 1,061.

Program Types

During each year from 1984 to 1988, a little more than one-fifth of the USMGs attended a program ranked as a top U.S. program and a little less than one-fourth attended a non-university-affiliated residency program. Only 6% of the Canadian graduates attended the top U.S. programs over all five years, and similarly small percentages attended non-university-affiliated programs. By far, the most typical program was a university program. This is because about three-fourths of the Canadian graduates attended residency programs in Canada, all of which are university or university-affiliated programs. Osteopathic graduates were most likely to be in university programs and university-affiliated programs in 1984 and 1986, but overall, 50% were in non-university-affiliated programs. Finally, 60% of all the FMGs trained in non-university-affiliated programs.

Overall Clinical Competence

Over the five years, USMGs had an average rating of overall clinical competence of 6.7 (SD = 1.2). The Canadian graduates generally received ratings that fell just into the "superior" range of 7–9. Their lowest average was 6.9 (SD = 0.9) in 1987 and the highest was 7.2 (SD = 0.9), achieved in 1985. Average ratings for the osteopathic graduates were more variable, ranging from 6.3 (SD = 1.2) to 6.8 (SD = 1.1). The ratings for FMGs averaged 6.3 (SD = 1.2).

Components of Competence

Candidates were also rated by their program directors on eight components of clinical competence. Within each of the four subgroups of candidates, the five annual average ratings for each component were similar. Three observations were apparent: (1) for all the subgroups, the highest ratings were for humanistic qualities and professional attitudes and behaviors, while the lowest tended to be for history-taking, medical knowledge, physical examination skills, and procedural skills; (2) for each component, the Canadians received the highest average rating, followed by the USMGs, the osteopathic graduates, and the FMGs; and (3) the ordering of components within each subgroup was nearly the same for the USMGs, the Canadians, and the FMGs, whereas the osteopathic graduates were rated comparatively lower on medical knowledge but higher on procedural skills.

Examination Performance

Examination scores are shown in Table 1. Because the reference group was made up almost entirely of USMGs who were first-time takers of the examination, it follows that the average score for that group is always close to 500. Rankings of the four groups of candidates showed a consistent pattern within each of the five years. The highest scores were always achieved by the Canadian graduates, ranging from 524 (SD = 98) in 1984 to 570 (SD = 76) in 1987. The osteopathic graduates achieved average scores that were about one-half a standard deviation, or 50 points, lower than the average USMG scores and one-half a standard deviation higher than the average FMG scores.

The consistent pattern of differences observed among the four groups in types of residency programs attended, program directors' ratings, and average scores led the authors to perform analyses to explore how these factors, in combination with type of medical school attended, predicted examination scores. Data were combined over all five years and submitted to a multiple regression. The dependent variable was examination scores. The independent variables were (1) type of medical school, coded as three dummy (0/1) variables for Canadian, osteopathic, and foreign undergraduate medical education; (2) type of residency program; and (3) overall rating of clinical competence. Together, these variables accounted for 23.6% of the variability in the observed examination scores. The unique portions attributable to each variable were 4.5% of foreign medical schools, 0.2% to Canadian medical schools, 0.1% to osteopathic medical schools, 2.3% to type of residency, and 8.5% to program directors' ratings. The amounts of variation shared by osteopathic and Canadian medical schools were 0.1% with type of residency program and 0.2% with the overall rating of competence.

Table 1

Mean Scores of Four Groups of Medical School Graduates Who First Took the American Board of Internal Medicine Certifying Examination between 1984 and 1988*

	1984		1985		1986		1987		1988		Total	
Group	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
USMGs	499	100	499	101	500	100	498	100	498	100	499	100
Canadian	524	98	540	86	531	96	570	76	552	82	545	88
Osteopathic	440	93	452	104	446	108	452	116	465	108	453	107
FMGs	384	137	385	132	401	135	420	127	430	125	405	132

*The groups were United States medical school graduates (USMGs), Canadian medical school graduates, graduates of osteopathic schools in the United States, and graduates of foreign medical schools (FMGs).

Scores are on a scale standardized to a reference group's mean of 500 with a standard deviation of 100.

Conclusions

These findings are informative on several accounts. In addition to the fact that descriptive data regarding numbers of graduates and average examination performances are provided for two sizable groups of candidates that were previously unstudied, questions are raised about the importance of the graduates' different educational backgrounds. Undergraduate Canadian training and a residency in a Canadian university program are generally thought, at least by the ABIM, to be equivalent in quality to training in the United States. Although the separate effects of each are confounded in this study, data from this small subgroup of candidates support this assumption, and, in fact, suggest that the joint Canadian effect may even be better than a similar U.S. experience. The findings could, however, reflect a sampling issue. While over 200 residents seek internal medicine certification in Canada each year, only about half as many choose to sit for the ABIM examination.[14] The self-selection into the final step of the certifying examination process in this country may be more responsible for their outstanding performance than any specific facet of training.

The repeated finding that the osteopathic graduates, on average, did less well than did USMGs might also reflect a sampling issue, albeit in a less obvious way. Among the many reasons students attend osteopathic medical schools are because they are drawn to primary care and because of the unique philosophy of an osteopathic undergraduate education. The coupling of these two interests undoubtedly leads to diverse and important educational experiences. It might be, however, that additional areas of instruction that are also relevant to ABIM examination performance are less emphasized. This general condition of "less exam preparedness" may be exacerbated if the preference for general medicine steers osteopathic graduates away from the more tertiary-care-oriented university hospitals and toward the more community-oriented training programs. Unfortunately, the choice of attending a non-university-affiliated residency is, on average, associated with poorer certifying examination performance. At the risk of sounding too negative, however, it should be emphasized that the osteopathic graduates did well. Although not reported in the results section, their overall pass rate was 69% (compared with 84% for the reference group). This suggests that osteopathic graduates are an untapped reservoir of talented physicians. Stepped-up recruiting efforts directed at this group would seem a prudent strategy, given the internal medicine residency positions unfilled in the Match in recent years.[15]

This work was supported by the American Board of Internal Medicine but does not necessarily reflect its views or opinions.

References

1. Norcini, J. J., Grosso, L. J., Shea, J. A., and Webster, G. D. The Relationship Between Features of Residency Training and ABIM Certifying Examination Performance. *J. Gen. Intern. Med.* **2**(1987):330-336.
2. Norcini, J. J., Shea, J. A., Webster, G. D., and Benson, J. A., Jr. Predictors of the Performance of Foreign Medical Graduates on the 1982 Certifying Examination in Internal Medicine. *JAMA* **256**(1986): 3367-3370.
3. Shea, J. A., Norcini, J. J., Day, S. C., Webster, G. D., and Benson, J. A., Jr. Performance of U.S. Citizen Caribbean Medical School Graduates on the American Board of Internal Medicine Certifying Examinations, 1984 to 1987. *Teach. Learn. Med.* **1**(1989):10-15.
4. Benson, J. A., Jr., Meskauskas, J. A., and Grosso, L. J. Performance of U.S. Citizen-Foreign Medical Graduates on Certifying Examinations in Internal Medicine. *Am. J. Med.* **71**(1981):270-273.
5. Meskauskas, J. A., Benson, J. A., Jr., and Hopkins, E. Performance of Graduates of Foreign Medical Schools on the Examinations of the American Board of Internal Medicine. *N. Engl. J. Med.* **297**(1977): 808-810.
6. Norcini, J. J., Swanson, D. B., Grosso, L. J., and Webster, G. D. Reliability, Validity and Efficiency of Multiple Choice Question and Patient Management Problem Item Formats in Assessment of Clinical Competence. *Med. Educ.* **19**(1985):238-247.
7. Norcini, J. J., Swanson, D. B., Grosso, L. J., Shea, J. A., and Webster, G. D. A Comparison of Knowledge, Synthesis, and Clinical Judgment Multiple-Choice Questions in the Assessment of Physician Competence. *Eval. Health Prof.* **7**(1984): 485-500.
8. Webster, G. D., Shea, J. A., Norcini, J. J., Grosso, L. J., and Swanson, D. B. Strategies in Comparison of Methods for Scoring Patient Management Problems. *Eval. Health Prof.* **11**(1988):231-248.
9. Norcini, J. J., Fletcher, S. W., Quimby, B. B., and Shea, J. A. Performance of Women Candidates on the American Board of Internal Medicine Certifying Examinations, 1973-1982. *Ann. Intern. Med.* **102**(1985): 115-118.
10. Appendix III, Medical Schools in Canada. *JAMA* **260**(1988):1157-1161.
11. Statistical Tables on the Osteopathic Profession. In *1988-89 Yearbook and Directory of Osteopathic Physicians.* Chicago, Illinois: American Osteopathic Association, 1988.
12. Crowley, A. E., and Etzel, S. I. Section III: Graduate Medical Education in the United States. *JAMA* **260**(1988):1093-1101.
13. *Policies and Procedures.* Philadelphia, Pennsylvania: American Board of Internal Medicine, July 1988.
14. *Annual Report and Performance Handbook: 1987.* Ottawa, Canada: Royal College of Physicians and Surgeons, 1988.
15. Aday, L. A., et al. National Study of Internal Medicine Manpower XII. The Future of Graduate Medical Education in Internal Medicine: What Do Program Directors Predict? *Arch. Intern. Med.* **148**(1988): 1509-1514.

A3273-8

OCTOBER 1990

ACADEMIC MEDICINE

[illegible]

A Perspective on Medical Expertise

Paradigms of Research

The Interactive Videodisc

Biomedical Nomenclature

Grading Systems for Clerkships

Selecting a Residency Program

Residents' Attitudes about Diabetes

Initial Encounter with a Cadaver

Using a Sociomatrix for Evaluation

Scientific Writing Courses

Students' Procedural Skills

National Policy Perspectives

VOLUME 65 • NUMBER TEN

OCTOBER 1990
VOLUME 65
NUMBER TEN

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

Special Article A Cognitive Perspective on Medical Expertise: Theory and Implications, *by* H. G. *Schmidt*, *Ph.D.*, G. R. *Norman*, *Ph.D.*, *and* H. P. A. *Boshuizen*, *Ph.D.* 611

Ideas for Medical Education Shifting Paradigms of Research in Medical Education, *by David* M. *Irby*, *Ph.D.* 622

Academic Computing in Medicine The Interactive Videodisc in Healthcare, *by Robert* F. *Rubeck*, *Ph.D.* 624

Developing the Videodisc for the Health Sciences: A Consortium Approach, *by Paula* N. O'Neill, *Ed.D.* 624

Medical Informatics Of Codes and Keywords: Standards for Biomedical Nomenclature, *by Daniel* R. *Masys*, M.D. 627

From the President 630

AAMC Focus 631

National Policy Perspectives 632

Research Reports
ARTICLES

A National Survey of Grading Systems Used in Medical Clerkships, *by Gregory* J. *Magarian*, M.D., *and Dennis* J. *Mazur*, M.D., *Ph.D.* 636

Factors Important to Students in Selecting a Residency Program, *by Albert* C. *Simmonds* IV, M.D., *James* M. *Robbins*, M.D., *Mark* R. *Brinker*, M.D., *Janet* C. *Rice*, *Ph.D.*, *and Morris* D. *Kerstein*, M.D. 640

SHORT PAPERS

Pediatrics Residents' Attitudes about Insulin-dependent Diabetes Mellitus and Children with Diabetes, *by Gary* M. *Ingersoll*, *Ph.D.*, *Roberta Ann Hibbard*, M.D., *Kathleen* K. *Kronz*, *Naomi* S. *Fineberg*, *Ph.D.*, *David* G. *Marrero*, *Ph.D.*, *and Michael* P. *Golden*, M.D. 643

Reactions of First-year Medical Students to Their Initial Encounter with a Cadaver in the Dissecting Room, *by* *David J. de L. Horne*, Ph.D., *John W. G. Tiller*, M.D., *Norman Eizenberg*, M.B.B.S., *Marika Tashevska, and Nola Biddle* 645

An Attempt to Motivate Internal Medicine Housestaff to Obtain Consent for Autopsies, *by Jaan Sidorov*, M.D. 647

A Comparison of the Emphasis Given to Nine Therapies in Residency Training and Later Practice of Psychiatrists, *by Ross E. Carter*, Ph.D. 649

Scientific Writing Courses for Pediatrics Fellows, *by Barbara G. Cox*, Ed.S., *Desmond A. Schatz*, M.D., *and* *L. H. S. Van Mierop*, M.D. 652

Using a Sociomatrix to Evaluate the Effectiveness of Small-group Teaching to Residents, *by William L. Toffler*, M.D., *Ann E. Sinclair*, M.S., *Marilyn S. Darr*, M.D., *Dean L. McGinty*, M.D., *Kate Commerford*, Ph.D., *and* *Rupert Goetz*, M.D. 654

The Procedural Skills of Medical Students: Expectations and Experiences, *by Vicki Kowlowitz*, Ph.D., *Peter Curtis*, M.D., *and Philip D. Sloane*, M.D. 656

Evaluating a Seminar Designed to Improve Psychiatry Skills of Family Medicine Residents, *by Jay H. Moss*, M.D. 658

BRIEF COMMUNICATIONS

The Presence of Student-based Peer Advising, Peer Tutoring, and Performance Evaluation Programs among U.S. Medical Schools, *by Maggie Moore-West*, Ph.D., *Sue Ann Hennessy*, *Philip W. Meilman*, Ph.D., *and* *Joseph F. O'Donnell*, 660; Identifying Personnel to Receive Continuing Medical Education about Appropriate Use of Blood Transfusions, *by Lawrence T. Goodnough*, M.D., *Alan L. Hull*, Ph.D., *Patricia Flately Brennan*, Ph.D., R.N., *Valentina A. Shah, and Barbara Martin*, 661; Hand Function in Selection of Orthopedics Residents, *by M. Mark Hoffer*, M.D., *and Shuchuan-Chen Hsu*, M.D., O.T.R., 661; A Medical Problem-solving Project for First-year Medical Students, *by Michael P. Rosenthal*, M.D., *Christopher Olbrich, and Edward H. McGehee*, M.D., 662; Applying Problem-based Learning Theory to the Clinical Clerkship, *by Theresa M. Romzick*, M.D., *and* *Robert W. Smith*, M.D., 662

Publishing Services: **Hanley & Belfus, Inc.**
210 South 13th Street, Philadelphia, PA 19107

OCTOBER 1990
VOLUME 65
NUMBER TEN

ACADEMIC MEDICINE

JOURNAL OF THE
ASSOCIATION OF
AMERICAN MEDICAL
COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
Edward J. Stemmler, M.D.
  Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
  St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
  Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
  Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
  University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
  University of Kentucky College of Medicine

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
  Georgetown University Hospital
Philip Anderson, M.D.
  University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
  Harvard Medical School
Eta Berner, Ed.D.
  University of Alabama School of Medicine
Ruth Bulger, Ph.D.
  National Academy of Sciences
Vincent A. Fulginiti, M.D.
  Tulane University School of Medicine
Nancy Gary, M.D.
  Albany Medical College
Beatrix Hamburg, M.D.
  Mt. Sinai Hospital
Murray M. Kappelman, M.D.
  The University of Maryland School of Medicine
Ralph W. Muller
  University of Chicago Hospitals
Herbert Pardes, M.D.
  College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
  University of Alabama School of Medicine
W. Vickery Stoughton
  Toronto General Hospital
Martha Vaughan, M.D.
  National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
David H. Cohen, Ph.D.
  Northwestern University

Chairman-Elect
William T. Butler, M.D.
  Baylor College of Medicine

Immediate Past Chairman
D. Kay Clawson, M.D.
  University of Kansas Medical Center School of Medicine

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
  George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Joe Dan Coulter, Ph.D.
  University of Iowa College of Medicine

Chairman, Council of Teaching Hospitals
Raymond G. Schultze, M.D.
  University of California, Los Angeles Medical Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00/1 year; $110/2 years; $155/3 years; $30.00/year for students; and $7.00 per single copy. All other locations—$70.00/1 year; $130/2 years; $185/3 years; and $8.00 per single copy.

Subscriptions received before June 30 will be honored for the calendar year (January through December). Issues from January up to receipt of first issue of new subscription will be mailed automatically.

Subscription requests received after June 30th will be honored for the fiscal year (July through June). Subscribers who want back issues for January through June of that calendar year must purchase them as individual back issues at $7.00 each and as available in stock.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.
Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036; (202) 828-0416.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

RESEARCH REPORTS

ARTICLES

A National Survey of Grading Systems Used in Medicine Clerkships

GREGORY J. MAGARIAN, M.D., and DENNIS J. MAZUR, M.D., Ph.D.

Abstract—To better understand the variety in U.S. medicine clerkship grading terminology, the number of grading levels, and the distribution of grades within each level, the authors surveyed medicine clerkship directors in the United States and Puerto Rico in 1986 and 1987. Completed questionnaires were returned from 101 of the 124 medical schools (81%). Descriptor grades were the most commonly used form of grading of medicine clerks, used in 68 of the 101 responding schools; letter grades were used by another 28 of the schools; and four schools used a numerical grading system. Although descriptor grades were most commonly used for grading medicine clerks, there was lack of consistency in their use between schools. The authors discuss the data related to letter grades and descriptor grades across this representative sample of U.S. medical schools. *Acad. Med.* 65(1990):636-639.

Clerkship grades have several major effects. First, grades provide feedback to students regarding their levels of performance. In many respects, grades provide incentives for students to give their best efforts when there is reason to believe that such efforts will be recognized. In this regard, the designation of terms to describe grades is probably not critical, since students are already aware of the various tiers in the grading system and, therefore, where their grades fall. Second, grades provide feedback to the department and school regarding how well students are performing. Finally, grades in specific clerkships may have a strong influence on students' competitiveness toward obtaining the housestaff positions of their choice.[1-3]

Of major concern to third-year clerks facing career decisions are the effects their grades have in marketing them as housestaff candidates. On the other hand, a medicine housestaff program director must also attempt to understand the meaning of grades from a particular institution and clerkship relative to grades from other schools when preparing rank order lists in choosing applicants for internships.

The purpose of this study was to survey medicine clerkships of the United States and Puerto Rico to determine the variability in grading terminology used, the number of grading levels used, and the distribution of grades within each level in each institution.

Method

Survey questions were sent to medical school chairpersons of 124 departments of medicine in the United States and Puerto Rico in 1986-1987. Each was requested to give the questionnaire to the medicine clerkship director or individual otherwise most responsible in coordinating the medicine clerkship. The questionnaire asked for the type of grading system used, offering four choices: (1) honors, pass, marginal, and fail; (2) honors, pass, or fail; (3) A, B, C, and F; or (4) other, in which case the clerkship director was asked to specify the grading system. Second, the clerkship director was asked to delineate the distribution of grades given for each level within the grading structure. The questionnaire also asked whether one or more grades to distinguish clinical skills and knowledge were recorded on a student's transcript. Those using a grading system without an intermediate level between the highest level and the mid-level were asked whether they felt their grading system hindered the above-average performer whose final evaluation grade otherwise appeared identical to the grade given to those who performed at lower levels but who yet received passing grades.

Results

Questionnaires were completed and returned from 101 of the 124 medical schools (81%). Grades based on descriptors such as honors-pass or outstanding-satisfactory were used by 68 schools. Letter grades were assigned by 28 schools, and numerical rankings by four schools. Information on

---

Dr. Magarian is director of the medicine clerkship and associate professor, and Dr. Mazur is associate director of the medicine clerkship and assistant professor; both in the Department of Medicine, Oregon Health Services University, Portland. Both authors are also staff physicians in the Section of General Internal Medicine, the Portland Veterans Administration Medical Center, Portland, Oregon.

Correspondence and requests for reprints should be addressed to Dr. Magarian, VA Medical Center (111-P), 3710 SW US Veterans Hospital Road, P.O. Box 1034, Portland, OR 97207.

Table 1

Number and Percentage of 28 Medical Schools Using Each Letter Grade, within Each Tenth-Percentile Grouping of the Clerks Given Each Grade, 1986–87*

Percentile Grouping of Clerks Given Grade	Number and Percentage of Schools Using Each Grade									
	A		B		C		D		F	
	No.	%	No.	%	No.	%	No.	%	No.	%
0–10	3	11.5	0	0	8	29.6	5	100	27	100
11–20	15	57.7	0	0	9	33.3	0	0	0	0
21–30	4	15.4	3	11.5	5	18.5	0	0	0	0
31–40	2	7.7	2	7.7	2	7.4	0	0	0	0
41–50	1	3.8	5	19.2	0	0	0	0	0	0
51–60	0	0	4	15.4	0	0	0	0	0	0
61–70	0	0	8	30.4	0	0	0	0	0	0
71–80	1	3.8	4	15.4	1	3.7	0	0	0	0
81–90	0	0	0	0	0	0	0	0	0	0
91–100	0	0	0	0	0	0	0	0	0	0
Total	26		26		25		5		27	
Unknown	2		2		2		0		0	
No. schools not giving grade	0		0		1		23		0	

*The authors surveyed medicine clerkship directors at 124 U.S. and Puerto Rican medical schools and received completed questionnaires from 101. Letter grades were used by 28 of the 101 responding schools, as analyzed in the table.

the use of numerical grades was not analyzed further.

Letter Grades

Letter grades were used by 28 of the 101 responding schools. When letter grades were used, their distributions of use where A's—19%, B's—56%, C's—19%, D's—4% (when used), and F's—2%. Only five (18%) of these 28 schools using letter grades indicated the use of the letter grade D. All indicated they had a failing-grade designation of F; however, ten (36%) indicated not recalling having given this grade. The highest failure rate for a class of medicine clerkships reported was 10% (two schools).

Table 1 analyzes letter-grade distributions within each tenth-percentile grouping of the clerks given each grade. Only two schools reported giving more than 37% of their medicine clerks A's. Similarly, only three of the 25 schools using letter grades from whom grade distribution data were obtained (12%) reported giving C's to more than 30% of their clerkship students. B's were clearly the predominant grade given. In only two of 25 schools were more C's given than B's. Both of these schools specifically noted that letter grades were subcategorized into pluses and minuses (for example, A+, A, A−).

Descriptor Grades

Descriptor grades were used by 68 schools. The distribution of grades indicative of various levels of performance is provided in Table 2.

Various terms were used to name the descriptor grades. Of the 68 schools using such grades, the highest level of grade was called (in descending order): honors, by 57 schools; outstanding, by four; high honors, two; pass, two; superior, two; and high pass, one. Grades at this level, but above "pass" (66 schools), were awarded to 14% of the students in the 62 schools from which sufficient data were obtained to make this determination. Among those indicating use of this grade designation, the lowest frequency of its use was 3% (two schools) and the highest was 30% (one school), with only 11% of such schools awarding this grade to more than 20% of a class.

Second-level descriptor grades between the highest awarded grade, such as honors, and a grade indicating acceptable performance but without distinction, were used by 26 (38%) of the 68 schools using descriptor grades. In order designated, they were high pass (16 schools), excellent (three), high satisfactory (three), near honors (two), above average (one), and above expected competence (one). The average percentage of students receiving a second-level grade was 32% in the 24 of 27 schools using such a grade for which sufficient information was available to determine frequency of use. In 63% of these 24 schools giving a second-level, intermediate grade, it was awarded to 30% or less of their student clerks. Only one school gave this grade to more than half of its students.

The grade given to the large number of students who performed adequately but without distinction in the 68 schools using descriptor grades was, in descending order: pass, by 55 schools; satisfactory, by seven; very good, two; good, one; acceptable, one; average, one, and expected competence, one. Two schools reported using a pass-fail grading system; therefore "pass" was the highest grade given to medicine clerks in these schools. Of these schools, the percentage of students receiving pass was 100% in one and 88% in the other. In one school, "high pass" was the highest grade awarded, whereas in two other schools, there was a grade

Table 2

Number and Percentage of 68 Medical Schools Using Descriptor Grades for Each Level of Performance, within Each Tenth-Percentile Grouping of the Clerks Given Each Type of Grade, 1986–87*

Percentile Grouping of Clerks Given Grade Type	Number and Percentage of Schools with Descriptor Grades for Each Level									
	Highest†		Intermediate‡		Midlevel		Marginal		Failure	
	No.	%	No.	%	No.	%	No.	%	No.	%
0–10	27	40	1	1	0	0	30	44	49	92
11–20	28	41	5	7	0	0	1	1		
21–30	7	10	9	13	1	1				
31–40			5	7	6	9				
41–50			3	4	8	12				
51–60			1	1	4	6				
61–70					9	13				
71–80					18	26				
81–90					15	22				
91–100					2	3				
Not given	2	3	40	59	0	0	31	46	13	19
Unknown	4	6	4	6	5	7	5	9	6	9

*The authors surveyed medicine clerkship directors at 124 U.S. and Puerto Rican medical schools and received completed questionnaires from 101. Grades indicating the various levels of performance shown in the table were used by 68 of the 101 responding schools.

†Highest = high honors, honors, outstanding, and superior. When a grade of "pass" or other grade of nondistinction was the highest available grade, it was not included as a highest-level grade.

‡Intermediate = high pass, excellent, high satisfactory, commendation, near honors, above average, above average competency.

above honors, designated "high honors."

When no intermediate grade was given between the highest awarded and an acceptable but nondistinctive performance, the descriptor—usually called "pass"—assigned to this level of performance was awarded 80% of the time. However, when there was an intermediate grade used, this midlevel grade was used less frequently, 56% of the time.

Grades for performances below levels that were clearly acceptable were called failures (used 1.7% of the time) or some variation of the term "marginal" (3.8%). It was not always clear whether marginal was considered a passing grade or indicated that some form of remediation was required.

Discussion

The 81% rate of responses to this survey from schools throughout the country suggests that these data represent overall trends in grading in U.S. medicine clerkships. This level of response and the written comments by responding clerkship directors also suggest they have a strong desire to know what their counterparts are doing.

At the time of this study, descriptor grades were the most commonly used form of grading of medicine clerks in U.S. medical schools, being used by 68 of the 101 responding schools. However, there was a lack of consistency in the use of descriptor grades between schools. Most used a system incorporating "honors" or a similar term for the highest level of performance, "pass" or a comparable term for acceptable but nondistinctive performance, and "fail" for unacceptable performance.

The highest-level descriptor grades, most commonly called "honors," were clearly grades of distinction in most medicine clerkships. Of the 68 schools using descriptor grades, the highest grade was given to an average of 14% of the students. Only 11% of the medical schools listing descriptor grades reported rewarding more than 20% of a class with this grade designation and, of these, none gave this level of grade to more then 30% (Table 2). Essentially equal numbers of medicine clerkships awarded this grade to 1–10% and 11–20% of their classes.

The authors have observed that within the large middle group of students receiving passing grades, but without distinction, there is a subset of clerks who perform much closer to the highest level just described than do the remaining students. Without a separate grade-recognition category, there is potential for such students to blend in with the larger number of students who do not truly perform at this level. Therefore, the large group of students who pass may include those students who perform at a near-honors level, as well as those who barely pass.

In this regard, 43% (29 of 68) of the schools using descriptor grades reported using an intermediate descriptor grade between the highest-level descriptor (such as "honors") and the larger middle level (such as "pass"). Of the seven terms reported to describe this level of performance, "high pass" (55%) was most commonly used. In these schools, this level of grade was awarded to an average of 28% of students. When there was no intermediate grade, the percentage of "honors" awarded increased only slightly to 20% from an overall frequency of 14%, but most who would have received an intermediate grade joined the larger middle pool, whose percentage increased to 80%. In other words, when an intermediate grade between an honors level and a middle level does not exist, students are more likely to receive middle-level grades than to have their grades upgraded to the next level, thereby losing any potential deserved recognition that an intermediate grade would have provided. Sixty-eight percent of the clerkship directors who had grading systems without intermediate grades between the highest level and the much larger middle group felt their students performing at this level were not being fairly recognized. Lack of recognition by a separate grade that distinguishes students from the large number below them may reduce a student's attractiveness as a potential house officer to program directors who use grades in clerkships of declared interest, among other factors, to assist in rank-ordering candidates for housestaff positions.

A middle-level grade (such as "pass") was given to more than half of the students in a class in approximately 90% of the schools using descriptor grades. The overall average number of medicine clerks receiving such a grade was 67.7%; however, as noted above, the percentage of its use is affected by whether a near-honors level of grade is used. The percentage of students receiving a middle-level grade decreased to 56% when an intermediate grade was awarded to a near-honors level of performance and increased to 80% when such an intermediate level did not exist. Only two schools awarded only "pass" or "fail" grades without delineation of other levels of performance.

Marginal and failing grades were infrequently given. Only 23 of the 68 schools using descriptor grades indicated use of a marginal grade. Various terms were used, including marginal, marginal pass, marginal fail, low pass, below expected competency, and below average. In this regard, it may even be confusing to one reviewing a transcript to know whether this grade is considered passing. Regardless, it was used only 3.8% of the time in those schools giving the grade, with only one school reporting as many as 10% of its students receiving this grade. "Failure" was given even less frequently (1.7%), with the highest number receiving this grade also being 10% in one school.

Letter grades were used by only 28 of the 100 responding schools, and only four schools used a numerical grading system. In schools using letter grades, A's were given in a frequency similar to that in effect for the highest descriptor grades, 18.9% versus 14.3%. Letter grades of A were subcategorized into A+ and A− in only three of 28 schools. Otherwise, for the most part, B, and not C, was usually comparable to the midlevel grade (passing). In only two schools did the number of C's awarded outnumber the number of B's. B's were given to 55.6% of students in schools using letter grades, compared with a passing midlevel descriptor grade given to 67.7% of students, or 56% when an intermediate, near-honors grade was used. In only five other schools were the numbers of C's even half the numbers of B's awarded. All of this suggests that, except in two of 28 schools using letter grades (7%), a letter grade of B is not synonymous with an intermediate grade between "honors" and "passing" in the descriptor grading system, but, in fact, is more comparable to a midlevel descriptor grade such as pass, satisfactory, or acceptable.

The D and F grades do appear to be comparably used as marginals or failures in the descriptor system and are used with equal frequency (D's, 4.1% versus marginals, 3.8%, and F's, 1.7 versus failures 1.7%). Combined, these account for only about 5% of students in medicine clerkships.

Recommendations

The authors believe that when the evaluation process allows sufficient discrimination of performance to identify an intermediate level of performance between the highest level of achievement and a passing grade without distinction, it should be awarded. It rewards the student by recognizing a job well done, as well as providing greater discrimination of performance for better-than-average students, who may otherwise be less easily recognized as they apply for housestaff positions.

The distributions of the same grade from one school to another can be quite different. Therefore, we believe that each school should be responsible for clearly indicating the distribution of grades in all course, but especially clerkships, such that the performances of students from different schools can be better compared. This would preferably be accomplished on a schools's transcript or as an attachment to the transcript. At the very least, information relating to courses within the student's field of interest (internship) should be included within the dean's letters.

When intermediate grades are not used, students are dependent on deans' or departmental letters to distinguish their levels of performance in clerkships related to their areas of specialty interest. Without a letter, such students, on medical school transcripts, appear to be inferior to identical students for whom grade designations exist.

In summary, the authors maintain that it would be valuable to have greater consistency in the grading systems of medical schools in the United States as well as greater delineation of grade distributions in all courses and clerkships on all medical school transcripts. Such measures would make possible a greater appreciation of a student's performance in any clerkship at any medical school.

References

1. Wagoner, N. E., Suriano, R., and Stoner, J. A. Factors Used by Program Directors to Select Residents. *J. Med. Educ.* **61**(1986):10–21.
2. Featherstone, H. J., and Ramsey, P. Analysis of Selection Criteria for Medical Residents. *Am. J. Med.* **75**(1983):687–690.
3. McNeil, M., Parker, R., Bankowitz, R., and Challinor, S. Correlates of Internal Medicine as a Residency Choice among Students at the University of Pittsburgh. *Clin. Res.* **37**(1989):811A.

hensive Review of Current Issues. *JAMA* **258**(1987):364-369.
8. *To Get Consent for a Post-What the Top Hospitals Say to Do. Res. Phys.* **13**(Aug 1967):56-65.
9. Schneiderman, H., and Gruhn, J. How—and Why—To Request an Autopsy. *Postgrad. Med.* **77**(1985):153-164.
10. Spiegel, J. S., Shapiro, M. F., Berman, B., and Greenfield, S. Changing Physician Test Ordering in a University Hospital. *Arch. Intern. Med.* **149**(1989):549-553.
11. Martin, A. R., Wolf, M. A., Thibodeau, L. A., Dzau, V., and Braunwald, E. A Trial of Two Strategies to Modify the Test Ordering Behavior of Medical Residents. *N. Engl. J. Med.* **303**(1980):1330-1336.
12. Korn, J. E., Schlossberg, L. A., and Rich, E. C. Improved Preventive Care Following an Intervention during an Ambulatory Care Rotation: Carryover to a Second Setting. *J. Gen. Intern. Med.* **3**(1988):156-160.
13. Cohen, S. J., W[illegible]erger, M., Hui, S. L., Tierney, W. M., and McDonald, C. J. The Impact of Reading on Physicians' Nonadherence to Recommended Standards of Medical Care. *Soc. Sci. Med.* **21**(1985): 909-914.
14. Gardner, R., Peskin, L., and Katz, J. L. The Physician, the Autopsy Request and the Consent Rate. *J. Med. Educ.* **48**(1973):636-644.
15. Katz, J. L., and Gardner, R. Request for Autopsy Consent. *N.Y. State J. Med.* **73**(1973):2591-2596.

A Comparison of the Emphasis Given to Nine Therapies in Residency Training and Later Practice of Psychiatrists

ROSS E. CARTER, Ph.D.

Abstract—This study compared the emphasis given to various therapies in six psychiatry residencies and by psychiatrists in their later practice. Eighty-seven psychiatrists and 29 residents were surveyed in 1987. The two groups reported that almost the same amounts of time of their respective training programs had been devoted to the nine therapies surveyed. Patterns of practice were observed, including the finding that those who were highly trained in a therapy practiced it more than did those who had less training. Both those who were highly trained and those who were not extended their practices to the use of therapies in which their training was not extensive. Problems in the study design and variables affecting this type of research are identified as guidance for future investigations. *Acad. Med.* 65(1990): 649-652.

The relation between residency training and later practice is difficult to assess. The physician's intervening experiences confound the effects of his or her early training, hindering efforts to show causal links. This difficulty is particularly great when assessing the impact of psychiatrists' training. Even though the field has been defined as diverse,[1,2] the main activity of psychiatrists is the practice of psychotherapy,[3] a process that is difficult to define in standard terms because it is carried out in a wide variety of ways by practitioners. Few data exist to show how the practice of psychiatry is affected by the amount of therapy training in residency, the type of residency program, or continuing education.

Method

Psychiatrists who graduated from six U.S. psychiatry residency programs between 1976 and 1987, and the psychiatry residents in the same programs, were contacted in 1987 through their programs' training directors.[4] A total of 294 psychiatrists and 116 residents were mailed a questionnaire; usable replies were returned by 87 (29.6%) of the psychiatrists and 29 (25%) of the residents.

The questionnaire requested demographic information and inquired about each physician's current practice and residency training. Each respondent was asked to rate the amounts of time during his or her residency training devoted to nine therapies (see Table 1) and the amounts of his or her current practice in which the same therapies were offered. The responses were obtained through a format based on a linear analog technique adapted from that of Karon and O'Grady,[5] which allows the expression of subjective impressions and results in linear measures.[6]

Results

The psychiatrists reported a mean length of practice of 4.06 years (SD = 10.2), extensive continuing education, and a mean of 14.8 hours of therapy per week (SD = 11.6) in treating patients within a typical age range of 18-50 years and with nearly equal numbers of inpatients and outpatients. The residents reported slightly fewer therapy hours spent in treating patients within the same age range; however, 84% of their patients were outpatients. Both groups re-

---

Dr. Carter is associate professor, Department of Psychiatry and Mental Health Sciences, Medical College of Wisconsin, Milwaukee.

Correspondence and requests for reprints should be sent to Dr. Carter, Department of Psychiatry and Mental Health Sciences, Medical College of Wisconsin, 3521 N. Prospect Avenue, Milwaukee, WI 53211.

Table 1

Significance Levels for Differences in Amounts of Time Spent in Practice of Each of Nine Therapies by Psychiatrists (P's) and Residents (R's) with the Highest and Lowest Levels of Training in Each Therapy, 1976–1987[*]

	Psychopharmacology		Somatic		Consultation/Liaison		Community Psychiatry		Marital		Family		Classic/Analytic		Group		Psychodynamic	
Therapy	P's	R's	P's	R's	P's	R's	P's	R's	P's	R's	P's	R's	P's	R's	P's	R's	P's	R's
Psychopharmacology	.001	.001	.01	—	.01	.04	—	—	—	—	—	—	—	—	—	—	—	—
Somatic	.04	.02	.001	.002	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Consultation/liaison	.03	.02	.001	.03	.01	.03	—	.05	—	.03	.004	—	.04	—	—	.02	—	.03
Community psychiatry	.001	.01	—	—	.02	.04	.003	.01	—	—	.007	—	—	—	.02	—	—	—
Marital	—	—	.02	—	.04	—	—	—	.001	.003	.001	.02	.006	—	.03	—	.001	.02
Family	—	—	—	.02	.006	.04	—	—	—	.002	.001	.02	.04	—	.001	—	.05	—
Classic analytic	—	—	.05	—	—	—	.01	—	.007	—	.006	.05	.004	.01	.02	—	—	.01
Group	—	—	—	—	—	—	.04	—	—	—	—	—	.01	—	.004	.01	—	—
Psychodynamic	—	—	—	—	.05	—	.001	—	—	—	—	—	—	.03	—	—	.004	.001

[*] The author sent questionnaires to 294 psychiatrists who graduated from six U.S. psychiatry residency programs between 1976 and 1987, and to the psychiatry residents in the same programs when the study was carried out in 1987. The questionnaire inquired about each physician's current practice and residency training in terms of the amounts of time devoted to nine therapies. Student's *t*-test was used to calculate the differences between the amount of time each therapy was practiced by the physicians with the lowest level of training and the amount of time each therapy was practiced by those with the highest level of training. Only the comparisons that showed a significant difference are shown in the table.

ported, in decreasing order, the same preferences for the following categories of therapy: individual, family, marital, and group.

For all nine therapies, the amounts of training and practice were highly correlated ($p < .001$ to .002) for both the psychiatrists and the residents. Both groups reported the greatest amount of their training and practice in psychopharmacology and psychodynamic therapy; the least in group, analytic, and marital therapy.

The author classed the responding physicians into three groups (low, middle, and high) according to the amount of training each reported in each therapy. Differences in the amounts of time devoted to the therapies in practice by those in the high and low groups were compared using Student's *t*-test. Only the comparisons that showed a significant difference are shown in Table 1.

For both the psychiatrists and the residents, those reporting more training in a particular therapy practiced that therapy more than did those with less training. Differences were also found between the most- and the least-trained groups in several related areas. Of the 39 significant differences between the most- and least-trained psychiatrists, 19 were also observed between the high and low groups of residents, as indicated in Table 1.

Amounts of training were compared according to residency programs using a one-way analysis of variance. The resulting F values were significant ($p < .001$) for all areas except consultation/liaison and marital therapy, indicating that the residencies differed in types and amounts of training. Further comparisons were made using the Newman-Kuels method.[7] These resulted in only one residency's being rated as comprehensive, with others rated as having high and low areas of strength and as being divided between biological and nonbiological therapies.

Discussion

The data suggest "specialist/generalist" practices organized not only around areas in which the physicians had had extensive training but also around other therapies in which they had experienced less training. For example, those with extensive training in somatic therapy practiced it as "specialists" but also practiced, as "generalists," other forms of therapy in which they were less trained. However, those who had had less training in somatic therapy but practiced it as generalists spent less time conducting this therapy than did the specialists in it. This relation was found consistently among both the psychiatrists and the residents who reported the highest and lowest levels of training in the therapies surveyed.

Consistency in residency education over time is suggested by the correspondence between the amounts of training and practice given to particular therapies by the responding physicians. The finding that both the residents and the psychiatrists practiced more in areas in which they were more extensively trained supports the truism that one does what one has learned to do. The strength of the relation between the amount of training and the practice time devoted to a therapy, however, indicates the primacy of residency education. It does not appear that the psychiatrists sought out continuing education in areas of weakness and then, with improvement, increased the time devoted to practicing those areas. Instead, the proportion of practice time devoted to a therapy appears to have been similar to the proportion of time devoted to it during the residency.

The fact that the psychiatrists did not limit their practices to therapies in which they were extensively trained suggests that other, nontraining, factors also influence practice. These are unknown but may include situational demands and the availability of alternate treatment sources.

Another dimension is that of the extent to which biological or nonbiological therapies are practiced and the difference in the extents of the spread of practice into related therapies. Those with extensive training in biological therapies tended to restrict their practices to biological therapies. In contrast, those with extensive training in nonbiological therapies practiced both biological and nonbiological ones. These findings suggest that psychiatrists and residents tend to specialize in either a biological therapy or a nonbiological therapy, with crossover occurring particularly for those practicing nonbiological therapy.

The data also support the generally held assumption that different residencies emphasize different areas of training. Of greater importance is that little or no training in certain areas was reported. One residency was rated extremely low in six areas of therapy; two were extremely low in four areas. Only one residency appeared to be "midstream" and to offer either a moderate or a high amount of training across all the areas of therapy surveyed. What problems this imbalance in training will cause in reaching the goal[8] of graduating psychiatrists with a variety of experience and with competencies in a number of areas remain to be seen.

In further research, attention should be given to the following approaches and variables, which the present study found were important. First, more accurate measures of the amount and type of continuing education and its effect on practice are needed. Second, the past and current emphasis given in a particular residency to therapy training should be assessed to account for change over time. Third, attention is needed regarding how residents' bias shapes what is learned and which residency is selected. Fourth, the availability of alternative sources of treatment and the types of patients referred should be assessed in order to gain more data on how psychiatrists extend their practices into different areas of therapy.

References

1. Langsley, D. G., and Hollender, M. H. The Definition of a Psychiatrist. *Am. J. Psychiatry* **139**(1982):81–85.
2. Langsley, D. G., and Yager, J. The Definition of a Psychiatrist: Eight Years Later. *Am. J. Psychiatry* **145**(1988):469–475.
3. Carter, R. Training General Psychiatry Residents in Family Therapy: Some Effects and Perspectives. *J. Psychiatr. Educ.* **9**(1985):159–167.
4. Robinowitz, C. B., Kay, J., and Taintor, Z., eds. *Directory of Psychiatry Residency Training Programs.* 3rd ed. Washington, D.C.: American Psychiatric Press, 1986.
5. Karon, B., and O'Grady, P. Quantified Judgments of Mental Health from the Rorschach, TAT and Clinical Status Interviews as Means of a Scaling Technique. *J. Consult. Clin. Psychol.* **34**(1970):229–235.
6. Remington, M., et al. Comparative Reliability of Categorical and Analog Rating Scales in the Assessment of Psychiatric Symptomology. *Psychol. Med.* **9**(1979):765–770.
7. Weiner, B. J. *Statistical Principles in Experimental Design.* New York: McGraw-Hill, 1962, pp. 85–89.
8. *Essentials of Accredited Residencies.* Accreditation Council for Graduate Medical Education. Chicago, Illinois: American Medical Association, 1987.